Dr. Valentin Gold

Sozialwissenschaftliche Fakultät

Platz der Göttinger Sieben 3
37073 Göttingen

valentin.gold@sowi.uni-goettingen.de
www.valentingold.de

Prof. Dr. Philipp Reuß, MJur (Oxford)

Professur für Bürgerliches Recht, Zivilprozessrecht, Internationales Privatrecht und Rechtsvergleichung

Platz der Göttinger Sieben 6
37073 Göttingen
T. +49 (0) 551 39-27683

p.reuss@uni-goettingen.de
www.reusz.eu

FORSCHUNGSPROJEKT
Künstliche Intelligenz und richterliche Entscheidungsfindung

Milestone 2 & 3: Abschlussbericht

12. November 2026

Kooperationspartner:
Niedersächsisches Justizministerium, Am Waterlooplatz 1, 30169 Hannover
Georg-August-Universität Göttingen, Wilhelmsplatz 1, 37073 Göttingen

Autoren:
Dr. Valentin Gold, Prof. Dr. Philipp Reuß, Jens Olheide, Moritz Kolbe, Niklas Middelbeck

Unter Mitarbeit von Dr. Joshua Kohler und Henrike Thomsen

INHALTSVERZEICHNIS

Inhaltsverzeichnis II

A. Vorwort 1

B. Kurzzusammenfassung der Forschungsergebnisse 3
I. Wesentliche Erkenntnisse aus technisch-ethischer Perspektive 3
II. Wesentliche Erkenntnisse aus rechtlicher Perspektive 4
III. Übergreifende Schlussfolgerungen 5

C. Analyse der technischen und rechtlichen Handlungsoptionen und Handlungsempfehlungen 7
I. Technische Handlungsoptionen und Handlungsempfehlungen 7
1. Einleitung 7
2. Grundlagen und Problemrahmen 9
a. Risikotypen: Idiosynkratische und systemische Risiken 9
b. Bestehende Risiken in der Rechtsprechung und richterlichen Entscheidungsfindung 9
c. Stakeholder: Betroffene Akteursgruppen 11
d. Opportunitätskosten der Nichtnutzung 13
3. Methoden der Risikoanalyse 16
a. Qualitative Risikoanalyse 17
b. Quantitative Risikoanalyse 18
c. Grenzen quantitativer Risikoanalyse 20
4. Evaluations- und Reflexionsrahmen für den KI-Einsatz 22
a. Zweck des Evaluationsrahmens 22
b. Risikodimensionen des KI-Einsatzes 22
c. Reflexions- und Evaluationsfragen für Anwender 24
aa. Funktionaler Charakter der Tätigkeit 24
bb. Reversibilität von Entscheidungsfolgen 26
cc. Erforderliche Fähigkeiten und Kenntnisse 26
dd. Struktur der Entscheidungsfindung 27
ee. Entscheidungsvarianz und Bias-Strukturen 28
ff. Algorithmisierbarkeit der Entscheidung 29
gg. Ziel des KI-Einsatzes 30
d. Menschliche Kontrolle als institutionelle Leistung 31
aa. Grenzen des Human-Oversight-Modells 32
bb. Vier Bedingungen wirksamer Aufsicht 33
cc. Meaningful Human Intervention 33
dd. Typische Risikopfade 34
e. Qualitative Risikoanalyse: Szenarien 34
aa. Szenario 1: Automation Bias durch epistemische Überbewertung 34
bb. Szenario 2: Institutionelles Handicapping effektiver Kontrolle 35
cc. Szenario 3: Schleichende Standardisierung von Entscheidungspraktiken 35
dd. Szenario 4: Fehlinterpretation der epistemischen Grundlage algorithmischer Ausgaben 36
f. Reflexions- und Evaluationsfragen für Systemdesigner 37
aa. Interface-Gestaltung (Steuerungsarchitektur) 37
bb. Verhaltenslenkung und Choice Architecture 38

cc. Funktionsweise und Delegationsgrad 38
dd. Eingriffsmöglichkeiten und menschliche Kontrolle 39
ee. Epistemische Transparenz 39
ff. Risikokommunikation 40
g. Reflexions- und Evaluationsfragen für Organisationen 40
aa. Erwartungs- und Rollenklarheit 40
bb. Kommunikationsstrukturen 41
cc. Umgang mit gewollter Nichtnutzung 41
dd. Qualifikation und gemeinsame Lernprozesse 42
ee. Monitoring, Fehlerkultur und kognitive Sicherungsmechanismen 42
h. Synthese: Zusammenführung der drei Perspektiven 42
aa. Konsequenzen für die Risikobewertung 43
bb. Offene Anschlüsse 43
i. Durchlaufendes Anwendungsbeispiel: KI-gestützte Strafzumessungsempfehlung 44
aa. Beschreibung des Anwendungsbeispiels 44
bb. Phase 1: Kontexteinordnung – Anwenderperspektive 45
(1) Kontextfaktoren im Detail 45
(2) Risikoprofil 46
(3) Zwischenfazit Anwenderperspektive 46
cc. Phase 2: Technische Gestaltung – Designerperspektive 47
(1) Interface-Gestaltung (C.I.4.f.aa) 47
(2) Verhaltenslenkung (C.I.4.f.bb) 47
(3) Delegationsgrad (C.I.4.f.cc) 47
(4) Eingriffsmöglichkeiten (C.I.4.f.dd) 48
(5) Epistemische Transparenz und Risikokommunikation (C.I.4.f.ee–C.I.4.f.ff) 48
(6) Zwischenfazit Designerperspektive 49
dd. Phase 3: Institutionelle Einbettung – Organisationale Perspektive 49
(1) Erwartungs- und Rollenklarheit (C.I.4.g.aa) 49
(2) Kommunikationsstrukturen (C.I.4.g.bb) 49
(3) Umgang mit Nichtnutzung (C.I.4.g.cc) 50
(4) Qualifikation (C.I.4.g.dd) 50
(5) Monitoring und kognitive Sicherungsmechanismen (C.I.4.g.ee) 50
(6) Zwischenfazit Organisationsperspektive 51
ee. Gesamtfazit: Bewertung des Anwendungsbeispiels 51
j. Durchlaufendes Anwendungsbeispiel: MAKI (Massenverfahrens-Assistenz mithilfe von KI) 52
aa. Beschreibung des Anwendungsbeispiels 52
bb. Phase 1: Kontexteinordnung – Anwenderperspektive 53
(1) Kontextfaktoren im Detail 54
(2) Risikoprofil 55
(3) Zwischenfazit Anwenderperspektive 56
cc. Phase 2: Technische Gestaltung – Designerperspektive 56
(1) Interface-Gestaltung (C.I.4.f.aa) 57
(2) Verhaltenslenkung (C.I.4.f.bb) 57
(3) Delegationsgrad (C.I.4.f.cc) 58
(4) Eingriffsmöglichkeiten (C.I.4.f.dd) 58
(5) Epistemische Transparenz und Risikokommunikation (C.I.4.f.ee–C.I.4.f.ff) 59
(6) Zwischenfazit Designerperspektive 59

dd. Phase 3: Institutionelle Einbettung – Organisationale Perspektive 59
(1) Erwartungs- und Rollenklarheit (C.I.4.g.aa) 60
(2) Kommunikationsstrukturen (C.I.4.g.bb) 60
(3) Umgang mit Nichtnutzung (C.I.4.g.cc) 61
(4) Qualifikation (C.I.4.g.dd) 61
(5) Monitoring und kognitive Sicherungsmechanismen (C.I.4.g.ee) 62
(6) Zwischenfazit Organisationsperspektive 63
ee. Gesamtfazit: Bewertung des Anwendungsbeispiels 63
5. Normative und institutionelle Konsequenzen 66
a. Verantwortungszuweisung im Kontext KI-gestützter richterlicher Entscheidungsfindung 67
aa. Das Responsibility-Gap-Problem 67
bb. Funktionale Verantwortungsverteilung 68
cc. Bedingte Verantwortung und institutionelle Vorbedingungen 69
dd. Rückbindung an das Anwendungsbeispiel 69
b. Effizienz-Risiko-Abwägungen 70
aa. Struktur zulässiger Trade-offs 70
bb. Unverhandelbare Schwellen 70
cc. Entscheidungskompetenz über Trade-offs 72
c. Moderatoren und Schutzmaßnahmen 72
aa. Systematik der Maßnahmen 72
bb. Technische Maßnahmen (Designerebene) 73
(1) Informationsdesign und Unsicherheitskommunikation 73
(2) Technische Maßnahmen gegen Konfabulation 73
(3) Funktionale Beschränkung 74
(4) Mindest-Benchmarks als Ausschlusskriterien 74
(5) Symmetrischer Begründungsaufwand 75
cc. Organisationale Maßnahmen 75
(1) Cognitive Forcing Mechanisms 75
(2) Monitoring mit Frühwarnschwellen 75
(3) Institutionell geschützte Nichtnutzung 76
(4) Schulung mit Schwerpunkt auf Systemgrenzen 76
dd. Verfahrensbezogene Maßnahmen 76
(1) Externe Audits 76
(2) Fehleranalyse als systemische Untersuchung 77
(3) Transparente Dokumentation der Systemnutzung 77
ee. Zuordnung zu Risikopfaden 77
ff. Verhältnismäßigkeit der Maßnahmen 78
6. Schlussbetrachtung 79
a. Ertrag der Analyse 79
b. Grenzen des Rahmens 79
c. Das Regulierungsdilemma: Laufen, um zu stehen 80
d. Konsequenzen für die Justiz 82
e. Ausblick 83
II. Rechtliche Handlungsoptionen und Handlungsempfehlungen 84
1. Einleitung 84
2. Menschliche Aufsicht und Kontrolle des Systems 86
a. Übergreifende rote Linien für jeden Einsatz sowie Handlungsleitlinien 86
aa. Übergreifende rote Linien 86
(1) Umfassende Einbindung menschlicher Richterinnen und Richter 86
(2) Ausschluss von Bias 87

bb. Handlungsoptionen und -empfehlungen 88
(1) Richterliche Aufsicht und Kontrolle bei Beschaffung, Betrieb, Überwachung und Entwicklung 88
(2) Ansiedelung der Entwicklung und des Betriebs des informationstechnischen Systems 93
(3) Überwachung der Arbeitsergebnisse; Umgang mit Bias und sonstigen Fehlermustern 97
(4) Ausschluss inhaltlicher Kontrolle richterlicher Dokumente durch Verwaltung oder Dritte 105
(5) Diskriminierungssensible Auswahl und Kontrolle der Trainingsdaten 105
cc. Zusammenfassung 107
b. Rote Linien für einen entscheidungs*ersetzenden* Einsatz sowie Handlungsleitlinien 109
aa. Rote Linien entscheidungs*ersetzende*r Einsatz 109
bb. Handlungsoptionen und -empfehlungen 110
(1) Identifikation hinreichend einfach gelagerter Fälle und (Teil-)Einsatzszenarien 111
(2) Verfahrensrechtliche Ausgestaltung des entscheidungsersetzenden Einsatzes 112
(3) Jederzeitige Aussteuerungs(möglichkeit) in komplexeren Fallkonstellationen 114
(4) Vermeidung systemseitiger Halluzinationen 115
cc. Zusammenfassung 116
c. Rote Linien für einen entscheidungsunterstützenden Einsatz sowie Handlungsleitlinien 117
aa. Rote Linien entscheidungsunterstützender Einsatz 117
bb. Handlungsoptionen und -empfehlungen 119
(1) Substantielle Kontroll- und Korrekturmöglichkeiten 119
(2) Kontrollmöglichkeiten und systemseitige Kontrollpflichten 123
(3) Sicherstellung von Kontrolle und Aufsicht im Hinblick auf psychologische Effekte 126
(4) Information über Funktionsweise und Limitierungen des Systems 129
cc. Zusammenfassung 131
3. Schutz informationstechnischer Systeme vor Störungen 134
a. Übergreifende rote Linien für jeden Einsatz 134
b. Handlungsoptionen und -empfehlungen 134
aa. Schutz vor internen oder externen Störungen; Sicherstellung der ordnungsgemäßen Funktion und des gesetzmäßigen Betriebs 134
bb. Etablierung eines Risikomanagementsystems 139
cc. Genauigkeit, Robustheit und Cybersicherheit 143
dd. Systemseitige oder sonstige Vorkehrungen gegen unzulässigen Einfluss Dritter auf Entwicklung, Gestaltung und Kontrolle des Systems 146
c. Zusammenfassung 149
4. Transparenz staatlichen Handelns und Information über den Einsatz von Künstlicher Intelligenz 151
a. Übergreifende rote Linien für jeden Einsatz 151
b. Handlungsoptionen und -empfehlungen 152
aa. Äußere Transparenz gegenüber Parteien und Öffentlichkeit 152
(1) Information über das „Ob“ und das „Wie“ des Einsatzes 153
(2) Transparenzpflichten nach Art. 50 ff. KI-VO 156
bb. Innere Transparenz gegenüber Richterschaft und Betreiber 173

(1) Umgang mit der sog. Black-Box-Problematik 174
(2) Transparenzanforderungen an Hochrisiko-KI-Systeme nach Art. 13 KI-VO, insbesondere Betriebsanleitung 175
cc. Überprüfbarkeit des Systems und Dokumentation 177
(1) Quellcode-Offenlegung 178
(2) Protokollierungspflichten durch Anbieter und Betreiber 180
c. Zusammenfassung 182
5. Datenschutz 185
a. Übergreifende rote Linien für jeden Einsatz 185
b. Handlungsoptionen und -empfehlungen 186
aa. Rechtsgrundlage der Datenverarbeitung 187
bb. Gewährleistung der Datensicherheit 190
cc. Eingabe personenbezogener Daten in Systeme, die nicht der Kontrolle der Justiz unterstehen 197
dd. Trainingsdaten anonymisiert oder nicht personenbezogen 198
ee. Sicherstellung von Betroffenenrechten 200
ff. Datenverarbeitung-Informationspflichten 202
c. Zusammenfassung 205
6. Notwendigkeit einer Rechtsgrundlage 208
a. Rote Linien für einen entscheidungsersetzenden Einsatz sowie Handlungsleitlinien 208
aa. Rote Linien entscheidungsersetzender Einsatz 208
bb. Handlungsoptionen und -empfehlungen 209
(1) „Ob“ und „Wie“ des entscheidungsersetzenden Einsatzes 209
(2) Zuständigkeitsregelungen 210
cc. Zusammenfassung 217
b. Rote Linien für einen entscheidungsunterstützenden Einsatz sowie Handlungsleitlinien 218
aa. Rote Linien entscheidungsunterstützender Einsatz 218
bb. Handlungsoptionen und -empfehlungen 218
cc. Zusammenfassung 221
7. Verfügbarkeit von Rechtsbehelfen 223
a. Übergreifende rote Linien für jeden Einsatz 223
b. Handlungsoptionen und -empfehlungen 223
c. Zusammenfassung 229
8. Barrierefreiheit des Rechtsschutzes 231
a. Übergreifende rote Linien für jeden Einsatz 231
b. Handlungsoptionen und -empfehlungen 231
c. Zusammenfassung 235
9. Sicherstellung der Freiwilligkeit der Systemnutzung durch Richterinnen und Richter 236
a. Rote Linien entscheidungsunterstützender Einsatz 236
b. Handlungsoptionen und -empfehlungen 236
c. Zusammenfassung 238
10. Gesamtzusammenfassung 239
a. Untersuchungsansatz und Leitprinzip der Handlungsempfehlungen 239
b. Übergreifende rote Linien, Handlungsoptionen und -empfehlungen 239
c. Entscheidungsunterstützender Einsatz 243
d. Entscheidungsersetzender Einsatz 245
e. Gesamtschau und Ausblick 248
III. Zusammenfassende Schlussfolgerungen 249

Literaturverzeichnis X

A. Vorwort

Das diesem Bericht zugrundeliegende Forschungsprojekt stellt die wissenschaftliche Begleitung eines Entwicklungs- und Forschungsvorhabens des Landes Niedersachsen dar, in welchem der KI-Einsatz unter realen Bedingungen entwickelt und getestet wird. Es untersucht die technisch-ethischen sowie die rechtlichen Rahmenbedingungen für den Einsatz von Künstlicher Intelligenz bei der richterlichen Entscheidungsfindung.[1]

Der nun vorliegende Abschlussbericht (Milestone 2 & 3) knüpft an den im Juli 2025 vorgelegten Bericht (Milestone 1) an und baut auf den dort bereits eingehend erarbeiteten roten Linien, den technisch-ethischen sowie rechtlichen Rahmenbedingungen des Einsatzes von Künstlicher Intelligenz bei der richterlichen Entscheidungsfindung im Zivilprozess, auf. Er diskutiert eingehend mögliche Handlungsoptionen, die einen rechtskonformen und technisch-ethisch vertretbaren Einsatz von informationstechnischen Systemen sicherstellen sollen. Hierbei werden jeweils konkrete Handlungsempfehlungen abgeleitet, die die praktische Entwicklung und Implementierung informationstechnischer Systeme unterstützen und leiten sollten.

Der Bericht richtet sich neben den konkret Projektbeteiligten vor allem an Wissenschaftlerinnen und Wissenschaftler sowie Praktikerinnen und Praktiker, die sich mit Fragen der Nutzung von Künstlicher Intelligenz in der Justiz befassen.

Die Erstellung des Berichts war nur möglich durch die engagierte und kompetente Zusammenarbeit eines interdisziplinären Teams. Besonders hervorzuheben ist die Herausforderung, eine gemeinsame Sprache zu finden, die technische, ethische und juristische Perspektiven in ein verständliches und produktives Begriffsgefüge übersetzt. Dass dies auch in diesem zweiten Bericht wieder gelungen ist, ist dem offenen Austausch und der großen Kooperationsbereitschaft aller Beteiligten zu verdanken. Unsere Arbeit hat auch in diesem Forschungsabschnitt wieder sehr stark davon profitiert, dass auch im Gesamtprojekt, d. h. mit Entwicklern, Testgerichten und der Justizverwaltung, ein regelmäßiger sowie intensiver Austausch möglich war und ist. Insbesondere ein gemeinsamer Ethikworkshop, der im Sommer 2026 an der Georg-August-Universität Göttingen veranstaltet worden ist, hat intensiven und bereichernden fachlichen Austausch ermöglicht.

Unser ausdrücklicher Dank gilt letztlich dem gesamten Team, das über die hier genannten Autorinnen und Autoren hinaus zur Entstehung des Berichts beigetragen hat. Besonders danken wir dem Team des Lehrstuhls Reuß für viele Stunden engagierter Mitarbeit, Durchsicht und Korrektur umfangreicher Berichtsteile.

[1] Die Arbeit an diesem Beitrag wurde zudem gefördert durch das Projekt TITAN – Technologische Intelligenz zur Transformation, Automatisierung und Nutzerorientierung des Justizsystems, gefördert von der Daimler und Benz Stiftung.

Namentlich zu danken ist: Herrn Jens Olheide, der sich als Co-Autor dieses Berichts mit sehr großem Engagement, fachlicher Begeisterung und herausragendem Einsatz und mit Blick gerade für die mühsamen und schwierigen Problemstellungen der Erarbeitung unseres grundlegenden Werks gewidmet hat. Zu erwähnen sind ebenfalls Herr Niklas Middelbeck, der ebenfalls als Co-Autor wertvolle Inhalte des Berichts beigesteuert hat sowie Frau Henrike Thomsen und Herr Dr. Joshua Kohler, die das Projekt bereits seit Beginn begleiten und bereichert haben. Herauszuheben ist ebenfalls Herr Moritz Kolbe, der als Co-Autor maßgebliche Tiefenbohrungen bei der Entwicklung des Reflexions- und Evaluationsrahmens vorgenommen hat.

Valentin Gold und Philipp Reuß Göttingen, im November 2026

B. Kurzzusammenfassung der Forschungsergebnisse

Das Forschungsprojekt „Künstliche Intelligenz und richterliche Entscheidungsfindung“ untersucht im Rahmen der wissenschaftlichen Begleitung eines Entwicklungs- und Forschungsvorhabens des Landes Niedersachsen die technisch-ethischen und rechtlichen Rahmenbedingungen für den Einsatz Künstlicher Intelligenz (KI) bei der richterlichen Entscheidungsfindung. Aufbauend auf dem ersten Forschungsbericht (Milestone 1), der die maßgeblichen roten Linien, die technisch-ethischen sowie rechtlichen Rahmenbedingungen des Einsatzes von KI bei der richterlichen Entscheidungsfindung, eingehend herausgearbeitet hat, widmet sich der vorliegende Abschlussbericht (Milestones 2 & 3) der Diskussion konkreter Handlungsoptionen und der Ableitung spezifischer Handlungsempfehlungen, die einen technisch-ethisch verantwortbaren sowie rechtskonformen Einsatz von informationstechnischen Systemen ermöglichen sollen. Hierzu werden die technisch-ethische und die rechtliche Perspektive eigenständig untersucht und in ihren wesentlichen Schlussfolgerungen abschließend zusammengeführt.

I. Wesentliche Erkenntnisse aus technisch-ethischer Perspektive

Die zentrale technische Erkenntnis lautet: Das Risiko liegt nicht im System als solchem, sondern in der Konfiguration aus Systemfunktion, Entscheidungssituation und institutioneller Einbettung. Ein und dasselbe System kann in einem Verfahren eine sinnvolle Entlastung sein und in einem anderen Verfahren die Grundlagen rechtsstaatlicher Entscheidungsfindung berühren. Maßstab der Bewertung ist dabei nicht der fehlerfreie Richter, sondern der Status quo: Auch menschliche Entscheidungen unterliegen kognitiven Herausforderungen. Der Verzicht auf ein nachweislich nützliches System ist daher ebenso begründungsbedürftig wie sein Einsatz. Ein Nutzungszwang folgt daraus nicht.
Kernstück des technischen Teils ist ein **Evaluations- und Reflexionsrahmen**, mit dem sich **ein konkreter Einsatz in drei Phasen prüfen** lässt: aus der **Perspektive der Anwenderinnen und Anwender (Kontexteinordnung), der Systemdesigner (technische Gestaltung) und der Organisation (institutionelle Einbettung)**. Der **Rahmen ist technologieneutral** angelegt und gilt für klassifizierende und prognostische Systeme ebenso wie für generative Sprachmodelle. Er ermöglicht eine begründete Einstufung des Risikoprofils, das insbesondere dann hoch ist, wenn ein empfehlendes System in einem hoch normativen und komplexen Feld eingesetzt wird oder Fehlentscheidungen faktisch irreversibel sind.
Eine menschliche Letztentscheidung allein beherrscht das Risiko nicht. Menschliche Beteiligung kann zur bloßen Zustimmungsgeste werden, die algorithmische Ergebnisse legitimiert, statt sie zu kontrollieren. **Wirksame Kontrolle setzt vier Bedingungen voraus: kausale Eingriffsmacht, epistemischen Zugang zu den Grundlagen des Systemergebnisses, Selbststeuerungsfähigkeit unter den tatsächlichen Arbeitsbedingungen und Anreize, die nicht auf die bequeme Übernahme lenken**. Fehlt eine dieser Bedingungen, entstehen vier typische Risikopfade: die epistemische Überbewertung von Systemergebnissen (Automation Bias), eine faktisch eingeschränkte Intervention, weil Abweichungen aufwändiger sind als die Übernahme, eine schleichende institutionelle Standardisierung und die Fehlinterpretation statistischer oder generierter Aussagen als geprüfte Aussagen über den Einzelfall.

An einer fiktiven Strafzumessungsempfehlung und am real erprobten System MAKI wird gezeigt, dass der **Rahmen nicht zu einem pauschalen Ja oder Nein führt**, sondern zu einer **Stufenlogik**: Eine Fallvergleichsanalyse kann vertretbar sein, eine Strafempfehlung nur unter strengen Bedingungen, eine als vorausgefüllter Wert angezeigte Empfehlung hingegen nicht. Bei MAKI zeigt sich zudem, dass bereits die Verdichtung eines Anhörungsprotokolls im Asylverfahren risikorelevant ist und dass geteilte Annotationen und Textbausteine eine ausdrückliche Regelung erfordern, damit die normativen Vorentscheidungen einer Person nicht ungeprüft in fremden Verfahren wirksam werden.
Aus der Analyse folgen **drei normative Konsequenzen**. Erstens ein **Modell bedingter Verantwortung**: Rechtlich bleibt die Verantwortung beim Gericht; wer Systeme einsetzt, ohne die Bedingungen wirksamer Aufsicht zu schaffen, trägt jedoch eine institutionelle Mitverantwortung. Zweitens **unverhandelbare Schwellen, an denen kein Effizienzgewinn einen Kompromiss rechtfertigt**: die faktische Verlagerung der Entscheidung auf das System, eine sachlich nicht gerechtfertigte Ungleichbehandlung anhand geschützter Merkmale und die Übernahme von Systemempfehlungen, die das Gericht nicht selbst begründen könnte. Drittens ein **Katalog technischer, organisatorischer und verfahrensbezogener Schutzmaßnahmen**. Besonders wirksam sind die funktionale Beschränkung auf informierende Funktionen, der Verzicht auf vorausgefüllte Standardwerte, ein gleicher Begründungsaufwand für Übernahme und Abweichung, eine eigene Einschätzung vor Einsicht in den Systemvorschlag, institutionell geschützte Nichtnutzung, ein nur aggregiertes Monitoring und externe Audits.

II. Wesentliche Erkenntnisse aus rechtlicher Perspektive

Die rechtliche Untersuchung clustert die aus dem Verfassungs-, Menschen-, Europa- und Zivilprozessrecht folgenden Einsatzgrenzen und entwickelt hieraus konkrete Handlungsoptionen sowie Handlungsempfehlungen im Sinne eines rechtlichen „Gold-Standards“. Wiederholt leitend ist dabei die bereits im ersten Bericht (Milestone 1) entwickelte dreigliedrige Funktionsmatrix: Je stärker ein System inhaltlich-normative Aufgaben übernimmt, je höher sein Automatisierungsgrad und je komplexer das Anwendungsfeld ist, desto höher sind die Anforderungen zur Einhaltung der rechtlichen Rahmenbedingungen.
Eine zentrale Empfehlung betrifft die **Sicherstellung richterlicher Aufsicht und Kontrolle**. Richterinnen und Richter müssen umfassend in Beschaffung, Betrieb und Überwachung der Systeme sowie – je nach Funktion – auch in deren Entwicklung eingebunden werden. Empfohlen wird die Etablierung **institutionell abgesicherter richterlicher Governance-Strukturen**, deren konkrete Ausgestaltung vom jeweiligen Einsatzszenario abhängt. Die **Ansiedelung von Entwicklung und Betrieb des informationstechnischen Systems bei der Judikative** wird dabei als vorzugswürdiger Weg angesehen, insbesondere bei Systemen mit inhaltlichem Einfluss auf die Entscheidungsfindung. Eine unzulässige Einflussnahme (der Exekutive oder privater) Dritter auf richterliche Entscheidungen ist auszuschließen.
Weitere wesentliche Handlungsempfehlungen betreffen die **Sicherstellung von Diskriminierungsfreiheit, Datenschutz, IT-Sicherheit und Transparenz**. Hierzu werden insbesondere ein kontinuierliches Risikomanagement, diskriminierungssensible Trainingsdaten, regelmäßige Systemüberprüfungen, technische und organisatorische Sicherheitsmaßnahmen sowie eine unaufgeforderte Information der Verfahrensbeteiligten sowie der Öffentlichkeit über das „Ob“ und „Wie“

des KI-Einsatzes empfohlen. Die Nachvollziehbarkeit der Systemergebnisse und ihre Überprüfbarkeit müssen gewährleistet sein.
Für den **entscheidungsunterstützenden Einsatz** bleibt die richterliche Letztverantwortung maßgeblich. Richterinnen und Richter müssen über substanzielle Kontroll- und Korrekturmöglichkeiten verfügen und über die Funktionsweise sowie Grenzen der Systeme informiert sein. Empfohlen werden insbesondere **herkunftssensitive Benutzeroberflächen** mit **direktem Zugriff auf Originalquellen**, **niedrigschwellige Korrekturmöglichkeiten** und **Maßnahmen zur Vermeidung** eines unkritischen Vertrauens in die Systemausgaben (Automation Bias). Eine Nutzung ist unter den derzeitigen Gegebenheiten freiwillig; auch faktische Nutzungszwänge sind auszuschließen. Für inhaltsbezogene Unterstützungsfunktionen wird zudem die Schaffung einer **bundeseinheitlichen gesetzlichen Grundlage** empfohlen.
Ein **entscheidungsersetzender Einsatz** kommt nach den Untersuchungsergebnissen gegenwärtig nur für hinreichend einfach gelagerte, rechtlich und tatsächlich überschaubare Fallkonstellationen in Betracht. Die Grenzen für dieses Einsatzszenario wahren derzeit nur **regelbasierte Systeme**, deren Einsatz insoweit **empfohlen** wird. Voraussetzung sind insbesondere eine hinreichend bestimmte gesetzliche Grundlage, klare Zuständigkeitsregelungen, die Vermeidung systemseitiger Fehler und jederzeitige Aussteuerungsmöglichkeiten an menschliche Richterinnen und Richter, wenn sich herausstellt, dass kein hinreichend einfach gelagerter Fall in diesem Sinne vorliegt. Der gerichtliche Rechtsschutz muss wirksam gewährleistet bleiben.

III. Übergreifende Schlussfolgerungen

Insgesamt hat sich gezeigt, dass der Einsatz von KI bei der richterlichen Entscheidungsfindung erhebliche Potenziale zur Entlastung der Justiz eröffnet, die technische Leistungsfähigkeit der Systeme allein jedoch weder die technisch-ethische Vertretbarkeit noch die rechtliche Zulässigkeit des Einsatzes gewährleistet. Erforderlich sind vielmehr eine konsequent an Systemfunktion, Entscheidungssituation und institutioneller Einbettung orientierte Risikobewertung, wirksame richterliche Kontrolle, geeignete technische und organisatorische Schutzmaßnahmen sowie eine tragfähige rechtliche und institutionelle Governance unter wirksamer Beteiligung der Richterschaft. Den rechtlichen Rahmen bilden die aus dem Verfassungs-, Menschen-, Europa- und Zivilprozessrecht abgeleiteten roten Linien. Er sollte durch die herausgearbeiteten Handlungsempfehlungen ausgefüllt werden, die sich nicht allein an der rechtlichen Zulässigkeit, sondern auch an Akzeptanz, Vertrauen und Entlastungswirkung orientieren. Übergreifend kehren dabei insbesondere die Ansiedelung der Entwicklung und des Betriebs der Systeme bei der Judikative als „Königsweg“, die ganzheitliche Adressierung des gesamten Lebenszyklus der informationstechnischen Systeme sowie die Präferenz für Konzentration und Bund/Länder-übergreifende Abstimmungen wieder. Idealiter zeitnah ausgefüllter, gesetzgeberischer Handlungsbedarf besteht insbesondere für inhaltsbezogene Unterstützungsfunktionen sowie für den entscheidungsersetzenden Einsatz. Für die Umsetzung des Einsatzes informationstechnischer Systeme bei der richterlichen Entscheidungsfindung sollten zunächst Fallgestaltungen, in denen eine Entscheidung auf eine Vielzahl von Fällen angewendet werden kann, prioritär anvisiert werden. Eine schrittweise Einführung kann hierbei eine früh einsetzende und sukzessiv ansteigende Entlastungswirkung bewirken. Die Funktionsfähigkeit des Rechtsstaats ist von kaum zu überschätzender Bedeutung. Informationstechnische Systeme bei der richterlichen

Entscheidungsfindung können ihren Beitrag hierzu leisten. Dies sollte genutzt werden. Die Wahrung richterlicher Unabhängigkeit, rechtsstaatlicher Verfahrensgarantien, der Diskriminierungsfreiheit und der Begründbarkeit gerichtlicher Entscheidungen bilden dabei die unverzichtbare Grundlage eines verantwortungsvollen KI-Einsatzes. Zahlreiche Befunde der Arbeitspakete des vorliegenden Forschungsprojekts sind ausdrücklich an den Stand gegenwärtiger Gegebenheiten geknüpft. Entsprechend sollte der Komplex der Künstlichen Intelligenz und richterlichen Entscheidungsfindung auch weiterhin fortlaufend eng begleitet werden.

C. Analyse der technischen und rechtlichen Handlungsoptionen und Handlungsempfehlungen

I. Technische Handlungsoptionen und Handlungsempfehlungen

1. Einleitung[2]

Künstliche Intelligenz hält Einzug in die richterliche Entscheidungsfindung, einen Bereich, der wie kaum ein anderer auf Begründbarkeit, Nachvollziehbarkeit und persönliche Verantwortung angewiesen ist. Empfehlende, klassifizierende, prognostische und zunehmend auch generativ arbeitende Systeme versprechen Entlastung bei wachsender Verfahrenslast, größere Konsistenz und eine Verringerung bekannter menschlicher Verzerrungen. Zugleich verschieben sie die Bedingungen, unter denen Entscheidungen zustande kommen, verantwortet und kontrolliert werden. Die zentrale Frage dieses Berichts lautet daher nicht, ob KI in der Justiz Risiken erzeugt, sondern unter welchen Bedingungen aus einem technisch funktionierenden System ein justizielles Risiko wird.
Ein erster Bericht dieses Projekts hat dafür das Fundament gelegt.[3] Das Forschungsprojekt ‚Künstliche Intelligenz und richterliche Entscheidungsfindung‘ wird vom Niedersächsischen Justizministerium gefördert, das auch die in Abschnitt C.I.4.j behandelte Software MAKI entwickelt.[4] Die Bewertung von MAKI beruht auf öffentlich zugänglichen Quellen sowie Auskünften der Projektverantwortlichen und ist als Anwendung des hier entwickelten Rahmens zu verstehen, nicht als Evaluation der Software. Der erste Bericht hat aus den einschlägigen Regelwerken und Ethikleitlinien Prinzipien für den KI-Einsatz in der Justiz herausgearbeitet, die justiziellen Anwendungsfelder klassifiziert und eine Typologie kritischer Einsatzkonstellationen entwickelt. Er endet mit der Forderung nach einer Bewertung entlang funktionaler, technischer und kontextueller Kriterien. Genau hier setzt der vorliegende Bericht an. Dabei verschiebt sich der Fokus von der Klassifikation von Systemen und Prinzipien hin zur Analyse der konkreten Bedingungen, unter denen ein System zum Risiko wird. Die Quintessenz, die dieser Bericht entwickelt, lässt sich in einem Satz zusammenfassen: Das Risiko liegt nicht im System als solchem, sondern in der Konfiguration aus Systemfunktion, Entscheidungssituation und institutioneller Einbettung. Ein und dasselbe KI-System kann in der einen Verfahrenskonstellation eine sinnvolle Unterstützung sein und in einer anderen die Grundlagen rechtsstaatlicher Entscheidungsfindung untergraben. Wer Risiken des KI-Einsatzes bewerten will, muss deshalb den Kontext analysieren, nicht allein die Technik.
Die Untersuchung greift hierbei auf Methoden der Risikoanalyse zurück und unterscheidet idiosynkratische von systemischen Risiken, verbindet qualitative Szenarienbildung mit den Möglichkeiten und Grenzen quantitativer Verfahren und bezieht die Opportunitätskosten der Nichtnutzung ein.

---

[2] Bei der Erstellung dieses Abschnitts wurde KI-gestützt gearbeitet: KI-Systeme wurden eingesetzt, um ergänzend zu recherchieren, die Argumentation an einzelnen Stellen zu schärfen und das Dokument gemeinsam mit den Autoren kollaborativ redaktionell zu überarbeiten. Sämtliche inhaltlichen Punkte – Thesen, Bewertungen und Schlussfolgerungen – stammen von den Autoren, wurden von ihnen geprüft und liegen allein in ihrer Verantwortung.
[3] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025.
[4] Die Autoren sind an der technischen Entwicklung der Software nicht beteiligt.

Das Kernstück ist ein Evaluations- und Reflexionsrahmen, der diese Perspektive operationalisiert. Er überführt Risikotypologie, Stakeholderanalyse und methodische Grundsätze in strukturierte Reflexionsfragen entlang dreier Perspektiven: jener der Anwender, die mit dem System arbeiten, jener der Systemdesigner, die seine Architektur prägen, und jener der Organisationen, die seinen Einsatz einbetten. Aus ihm folgen die normativen Konsequenzen des Berichts: eine an die tatsächliche Aufsichtsfähigkeit gebundene Verantwortungsverteilung, die Bestimmung unverhandelbarer Schwellen, an denen kein Effizienzgewinn einen Trade-off rechtfertigt, und ein systematischer Katalog technischer, organisationaler und verfahrensbezogener Schutzmaßnahmen.
Der Bericht richtet sich vorrangig an Leserinnen und Leser mit juristischem Hintergrund und konzentriert sich auf Risiken, die aus der unmittelbaren Interaktion zwischen Mensch und System im Entscheidungsprozess hervorgehen. Der Rahmen ist dabei bewusst technologieneutral angelegt: Die entwickelten Risikodimensionen und Aufsichtsbedingungen gelten unabhängig davon, ob ein System klassifiziert, prognostiziert oder generiert. Offen bleibt, ob sich diese Prinzipien auch auf Systeme übertragen lassen, deren Architektur nicht mehr auf einem einzelnen, prüfbaren Output beruht, sondern selbständig verkettete Handlungsschritte umfasst (vgl. C.I.6.c). Der Bericht gibt den Stand von Oktober 2026 wieder; dies betrifft insbesondere die Angaben zur KI-Verordnung, zum Prüfstandard A5 des BSI und zum Erprobungsstand der diskutierten juristischen Software.

2. Grundlagen und Problemrahmen

a. Risikotypen: Idiosynkratische und systemische Risiken

Für die Bewertung von KI-Systemen in der richterlichen Entscheidungsfindung ist ein präziser Risikobegriff zentral. In der interdisziplinären Risikoforschung wird Risiko als Möglichkeit nachteiliger Folgen unter Bedingungen der Unsicherheit verstanden.[5] Dabei lassen sich je nach Art der verfügbaren Informationen qualitativ verschiedene Formen von Unsicherheit unterscheiden. Zu differenzieren ist zwischen kalkulierbarem Risiko, also unsicheren Ereignissen mit zumindest modellierbaren Eintrittswahrscheinlichkeiten, und grundlegender Unsicherheit im Sinne Knights, bei der Wahrscheinlichkeiten oder relevante Ereignisqualitäten nicht verlässlich bestimmbar sind. Gerade bei datengetriebenen, adaptiven Systemen ist diese Differenz praktisch bedeutsam, da nicht alle Fehlermodi ex ante quantifizierbar sind.
Für eine strukturierte Risikobewertung ist darüber hinaus entscheidend, auf welcher Ebene sich ein Risiko manifestiert und welche Reichweite seine Folgen haben. Auf dieser Grundlage ist zwischen idiosynkratischen und systemischen Risiken zu differenzieren: **Idiosynkratische Risiken** betreffen einzelne Entscheidungen, Akteure oder Verfahrensschritte. Ihre Auswirkungen bleiben lokal begrenzt und sind typischerweise isolierbar sowie gezielt adressierbar. In Anlehnung an das NIST AI Risk Management Framework[6] entsprechen sie lokalen Risiken mit begrenztem Wirkungskreis. **Systemische Risiken** hingegen entstehen aus strukturellen Interdependenzen innerhalb komplexer, vernetzter Systeme.[7] Charakteristisch ist, dass Störungen einzelner Komponenten über Kopplungen, Standardisierung oder Skalierung zu Kaskadeneffekten führen können („contagion“). Das Risiko liegt hier weniger im einzelnen Fehler als in der Struktur des Systems und seiner Anfälligkeit für kollektive Fehlfunktionen. Der NIST-Rahmen beschreibt solche Risiken als übergreifend, mehrere Stakeholdergruppen betreffend und geprägt durch Emergenz, Persistenz sowie schwer steuerbare Wechselwirkungen.
Für den Einsatz von KI in der Justiz hat diese Unterscheidung unmittelbare Konsequenzen. Ein einzelner fehlerhafter Output kann idiosynkratisch bleiben, d. h. folgenreich für den Einzelfall, aber strukturell eingrenzbar. Wird ein System jedoch breit implementiert, standardisiert trainiert oder normativ prägend für Entscheidungsroutinen, können lokale Fehlannahmen strukturelle Verzerrungen erzeugen, die sich über viele Verfahren hinweg fortpflanzen. Mit zunehmender Integration verschiebt sich das Risikoprofil daher vom Einzelfallproblem zur potenziellen Systemwirkung. Eine solche Verschiebung erfordert qualitativ andere Bewertungsmaßstäbe und Kontrollmechanismen.

b. Bestehende Risiken in der Rechtsprechung und richterlichen Entscheidungsfindung

Die Unterscheidung zwischen idiosynkratischen und systemischen Risiken dient nicht allein als analytisches Werkzeug für die Bewertung von KI-Systemen, sie gilt ebenso

[5] *Society for Risk Analysis*, Society for Risk Analysis Glossary, 2018, abrufbar unter https://www.sra.org/risk-analysis-introduction/risk-analysis-glossary/ (Abruf am 25.10.2026, 14:20 Uhr).
[6] *National Institute of Standards and Technology (NIST)*, Artificial Intelligence Risk Management Framework (AI RMF 1.0), 2023, abrufbar unter https://doi.org/10.6028/NIST.AI.100-1 (Abruf am 25.10.2026, 14:32 Uhr).
[7] *Haldane/May*, Systemic risk in banking ecosystems, in: Nature 469 (2011), S. 351.

für den Status quo richterlicher Entscheidungsfindung. Eine Bewertung von KI setzt voraus, dass dieser Ausgangszustand realistisch eingeschätzt wird.

Auf der idiosynkratischen Ebene ist sie durch fehleranfällige Einzelfallentscheidungen, inkonsistente Normanwendung oder situativ verzerrte Wahrnehmung gekennzeichnet. Unter bestimmten institutionellen Bedingungen, wie etwa bei hoher Arbeitslast, unter implementierten und gelebten Arbeitsroutinen oder durch strukturelle Anreizverzerrungen, können sich daraus auch systematische Effekte entwickeln. Diese Risiken sind damit nicht kategorial verschieden von jenen, die für KI-Systeme diskutiert werden. Sie folgen derselben Logik, nur unter anderen technischen und institutionellen Bedingungen.

Teilweise sind diese Risiken institutionell adressiert. Instanzenzüge, Begründungspflichten, Kollegialentscheidungen und Verfahrensgarantien fungieren als Kontrollmechanismen, die Fehlentwicklungen begrenzen sollen. Ihre Reichweite ist jedoch begrenzt und empirische Befunde zeigen, dass kognitive Verzerrungen auch unter diesen Bedingungen wirksam bleiben.

Die empirische Entscheidungsforschung liefert dafür konsistente Belege. Edmond und Martire[8] zeigen, dass auch professionelle Entscheidungsträger automatischen Heuristiken und Wahrnehmungsverzerrungen ausgesetzt sind. Bystranowski et al.[9] belegen in einer Metaanalyse signifikante Ankereffekte in rechtlichen Entscheidungen; ob juristische Expertise den Effekt abschwächt, ist bislang nur vorläufig untersucht. Experimentelle Befunde zeigen zudem, dass Richterinnen und Richter unzulässige Informationen auch bei bewusster Anstrengung nur eingeschränkt aus ihrer Entscheidung heraushalten können.[10] Feldstudien deuten darüber hinaus auf den Einfluss situativer Faktoren wie Ermüdung hin,[11] wobei das Ausmaß dieser Effekte methodisch umstritten ist.[12]

Die Schlussfolgerung, dass richterliche Entscheidungen per se defizitär wären, kann aus diesen Befunden jedoch nicht gezogen werden. Vielmehr sind sie als Formen menschlicher Entscheidungsprozesse zu analysieren, die sowohl idiosynkratische Fehlerquellen als auch potenzielle strukturelle Verzerrungen aufweisen können. Damit unterliegen sie denselben analytischen Kategorien wie KI-Systeme.

Damit wird deutlich, dass Effizienz- oder Qualitätsversprechen von KI-Systemen sich nur sinnvoll beurteilen lassen, wenn der Vergleichsmaßstab, die menschliche Entscheidungsfindung, empirisch informiert ist. Normative Argumentationen dürfen daher nicht auf impliziten Tatsachenannahmen beruhen, die empirisch ungesichert sind. Eine verantwortbare Analyse des KI-Einsatzes erfordert die kontrollierte Einbeziehung empirischer Befunde in die rechtswissenschaftliche Argumentation.

Dabei bleibt jedoch zu beachten, dass die empirische Forschung selbst fallibel und meist nicht einheitlich ist. Sie unterliegt statistischer Unsicherheit, Modellannahmen, Kontextabhängigkeit und Reproduzierbarkeitsproblemen. Handlungsempfehlungen dürfen daher empirische Befunde weder ignorieren noch unkritisch verabsolutieren. Erforderlich ist ein reflektierter Umgang mit ihrer Reichweite, ihren Voraussetzungen und ihren Grenzen.

---

[8] *Edmond/Martire*, Just Cognition: Scientific Research on Bias and Some Implications for Legal Procedure and Decision-Making, The Modern Law Review 82 (2019), 633.

[9] *Bystranowski/Janik/Próchnicki/Skórska*, Anchoring Effect in Legal Decision-Making: A Meta-Analysis, Law and Human Behavior 45 (2021), 1.

[10] *Wistrich/Guthrie/Rachlinski*, Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding, University of Pennsylvania Law Review 153 (2005), 1251.

[11] *Danziger/Levav/Avnaim-Pesso*, Extraneous factors in judicial decisions, in: Proceedings of the National Academy of Sciences 108 (2011), S. 6889.

[12] *Weinshall-Margel/Shapard*, Overlooked factors in the analysis of parole decisions, Proceedings of the National Academy of Sciences 108 (2011), E833.

Für die Entwicklung späterer Evaluations- und Implementationskriterien folgt daraus einerseits, dass der Vergleich mit dem Status quo empirisch informiert sein muss, andererseits, dass sowohl die empirischen Annahmen als auch die daraus abgeleiteten Empfehlungen selbst überprüfbar und revisibel bleiben müssen.

c. Stakeholder: Betroffene Akteursgruppen

Da die Risiken in der richterlichen Entscheidungsfindung nicht isoliert auftreten, sondern stets in institutionelle und soziale Kontexte eingebettet sind, setzt eine systematische Risikoanalyse nicht nur eine Klassifikation von Risikotypen voraus, sondern auch eine Identifikation der relevanten Akteursgruppen. Richterliche Entscheidungen entfalten Wirkungen auf unterschiedlichen Ebenen und betreffen verschiedene Akteure in unterschiedlicher Intensität. Mit der Implementierung und Nutzung von KI-Systemen erweitert sich dieser Kreis potenziell betroffener Akteure, da neue technische und institutionelle Vermittlungsebenen hinzutreten. In diesem Zusammenhang ist zwischen direkt und indirekt betroffenen Akteuren sowie zwischen unterschiedlichen Formen ihrer Einbindung in den Entscheidungsprozess zu unterscheiden.
Dong et al.[13] zeigen etwa für KI-gestützte Entscheidungen über Sozialleistungen, dass unmittelbar Betroffene Genauigkeitsverluste zugunsten schnellerer Entscheidungen deutlich weniger akzeptieren als die allgemeine Bevölkerung. Zugleich überschätzen Nicht-Betroffene systematisch die Bereitschaft der Betroffenen, solche Abwägungen hinzunehmen. Überträgt man diesen Befund auf den Justizkontext, ist diese Wahrnehmungsasymmetrie von besonderer Relevanz: Richterliche Entscheidungsprozesse sind in ein dichtes Netzwerk institutioneller und gesellschaftlicher Akteure eingebettet, deren Perspektiven und Interessen erheblich divergieren können.
Vor diesem Hintergrund lassen sich drei Akteursebenen unterscheiden. Die **primäre Ebene umfasst unmittelbar Betroffene**. Hierzu zählen insbesondere Verfahrensbeteiligte sowie Richterinnen und Richter selbst, ferner Rechtspflegerinnen und Rechtspfleger, soweit ihnen richterliche Aufgaben übertragen sind, die sie sachlich unabhängig wahrnehmen. Entscheidungen entfalten für diese Akteure unmittelbare rechtliche und teilweise existenzielle Konsequenzen. Die **sekundäre Ebene bilden institutionell Beteiligte** wie Gerichte, Justizverwaltungen, Geschäftsstellen sowie unterstützende Berufsgruppen, die organisatorisch und strukturell in die Entscheidungsprozesse eingebunden sind und Verantwortung für deren institutionelle Umsetzung tragen. Die **tertiäre Ebene schließlich umfasst die gesellschaftliche Öffentlichkeit**: Zivilgesellschaft, Medien, Rechtswissenschaft und politische Akteure, die mittelbar betroffen sind, insofern gerichtliche Entscheidungen Erwartungen an Rechtsstaatlichkeit, Fairness und Legitimität prägen.
Die drei Ebenen ordnen Akteure nach dem Grad ihrer Betroffenheit, nicht nach ihrer Rolle im Umgang mit einem KI-System. Quer dazu liegt die Unterscheidung zwischen Anwendern, Systemdesignern und Organisationen, die dem Evaluationsrahmen in Kapitel C.I.4 zugrunde liegt (Anwender C.I.4.c, Systemdesigner C.I.4.f, Organisation C.I.4.g). Anwender sind diejenigen, die mit einem System im konkreten Entscheidungsprozess arbeiten. Das sind in erster Linie Richterinnen und Richter sowie Rechtspflegerinnen und Rechtspfleger, die als Entscheidende zugleich der primären Ebene angehören. Es können aber auch Mitarbeitende der Geschäftsstellen sein, die nach dieser Systematik der sekundären Ebene zuzurechnen sind. Für die

[13] *Dong/Bonnefon/Rahwan*, Heterogeneous preferences and asymmetric insights for AI use among welfare claimants and non-claimants, Nature Communications 16 (2025), Art. 6973.

Risikoanalyse ist diese Überkreuzung bedeutsam: Risiken, die in der Interaktion eines Anwenders mit dem System entstehen, wirken sich unabhängig von dessen eigener Ebene regelmäßig auf der primären Ebene aus, also bei den Verfahrensbeteiligten.
Zwischen diesen Ebenen bestehen regelmäßig divergierende Risiko- und Effizienzbewertungen. Für die Genauigkeit ist belegt, dass unmittelbar Betroffene Fehler algorithmischer Entscheidungen deutlich weniger hinzunehmen bereit sind als weniger stark involvierte Akteure.[14] Dass dies in gleicher Weise für Anforderungen an Transparenz und Fairness gilt, ist plausibel, aber empirisch nicht gesichert. Die Forschung zeigt vielmehr, dass die Fairnesswahrnehmung Betroffener stark davon abhängt, ob eine Entscheidung zu ihren Gunsten ausfällt.[15] Ihre Perspektive ist daher bei der Bewertung eines Systems unverzichtbar, als alleiniger Maßstab aber ungeeignet. Maßnahmen zur Risikoreduktion zugunsten einer Ebene können daher andere Ebenen unverhältnismäßig belasten. Diese Spannung muss bei der Entwicklung von Evaluationskriterien explizit berücksichtigt werden.
Mit der Einführung von KI-Systemen verändert sich diese Stakeholderlandschaft strukturell. Eine pauschale Betrachtung von „KI“ greift dabei zu kurz, da Risiken stets von der konkreten Funktion des Systems, seinem Einsatzkontext und seiner institutionellen Einbettung abhängen. Gleichwohl lassen sich auf abstrakter Ebene mehrere Verschiebungen identifizieren, die für die Risikobewertung zentral sind.
Erstens treten mit Entwicklern und Betreibern **neue Akteursgruppen** hinzu, die auch ohne formale Verfahrensbeteiligung Entscheidungsstrukturen durch Modellarchitektur, Trainingsdaten, Parametrisierung, Updates und Interface-Design beeinflussen. Da diese Akteure häufig privatwirtschaftlich organisiert sind, entstehen Governance-Herausforderungen hinsichtlich Transparenz, Auditierbarkeit und Verantwortungszuschreibung, die mit justizrechtlichen Instrumenten adressiert werden müssen.
Zweitens **transformieren KI-Systeme die Entscheidungsprozesse** selbst. Als Empfehlungssysteme, Filter oder Default-Optionen strukturieren sie Aufmerksamkeit, Hypothesenbildung und Beweiswürdigung. Selbst wenn die formale Letztentscheidung beim Gericht verbleibt, kann sich faktisch eine Begründungslast bei Abweichung von Systemempfehlungen etablieren. Dadurch verändern sich sowohl kognitive Entscheidungsprozesse auf der primären Ebene als auch institutionelle Erwartungsstrukturen hinsichtlich Effizienz, Konsistenz und Fehlertoleranz.
Die Interaktion zwischen Mensch und System wird damit zu einem eigenständigen Risikofaktor. Diese Entwicklung ist unabhängig von der tatsächlichen Systemqualität aus rechtsstaatlichen Gründen problematisch: Richterliche Unabhängigkeit setzt voraus, dass Abweichung vom technischen Default institutionell nicht sanktioniert wird, weil sonst Verantwortungszuschreibung systematisch diffundiert.
Drittens **verstärken breit eingesetzte, standardisierte Systeme das Potenzial für systematische Kaskadeneffekte.** Analysen vernetzter Systeme zeigen ein sogenanntes „robust-yet-fragile“-Muster[16]: Hohe Interkonnektivität kann kurzfristig Stabilität erzeugen, bei hinreichend großen Störungen jedoch systemweite Fehlfunktionen verstärken. Übertragen auf die Justiz kann extensive Standardisierung die natürliche Diversität richterlicher Praxis reduzieren. Kurzfristige Varianzreduktion

[14] *Dong/Bonnefon/Rahwan*, Heterogeneous preferences and asymmetric insights for AI use among welfare claimants and non-claimants, Nature Communications 16 (2025), Art. 6973.
[15] *Wang/Harper/Zhu*, Factors Influencing Perceived Fairness in Algorithmic Decision-Making: Algorithm Outcomes, Development Procedures, and Individual Differences, Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, 1.
[16] *Haldane/May*, Systemic risk in banking ecosystems, in: Nature 469 (2011), 351.

kann dabei langfristig mit systematischen Verzerrungen erkauft werden, die sich aufgrund der institutionellen Kopplung über viele Verfahren hinweg ausbreiten.
Viertens entstehen durch den Betrieb von KI-Systemen in komplexen technischen Infrastrukturen eigenständige **infrastrukturelle und datensicherheitsbezogene Risiken**. Zentralisierung, Plattformabhängigkeit und Vernetzung erhöhen die Anfälligkeit für Cyberangriffe, Systemausfälle oder Datenlecks. Diese Kategorie umfasst nicht nur klassische IT-Sicherheitsrisiken, sondern KI-spezifische Angriffsvektoren: Data Poisoning (gezielte Manipulation von Trainingsdaten), Model Extraction (Rekonstruktion eines trainierten Modells durch systematische Anfragen), Model Evasion (gezielte Umgehung von Modellentscheidungen durch adversariale Eingaben) sowie Membership Inference (Rückschluss auf einzelne Trainingsdatensätze und damit potenziell auf Verfahrensbeteiligte). Bei generativ arbeitenden Systemkomponenten tritt als weiterer Angriffsvektor Prompt-Injection hinzu: gezielt platzierte, für Menschen oft nicht wahrnehmbare Anweisungen – etwa unsichtbar in Schriftsätzen platziert –, die ein System zu einem von der eigentlichen Aufgabe abweichenden Verhalten verleiten sollen. Es sind inzwischen einige Fälle dokumentiert, in denen leistungsfähige Sprachmodelle in agentischer Form bereits eingesetzt wurden, um Sicherheitslücken in fremden IT-Systemen weitgehend selbständig aufzuspüren und auszunutzen.[17] Weitere Berichte zeigen zudem erste Schadsoftware, die während ihrer Ausführung auf generative Modelle zugreift[18], auch wenn KI bislang überwiegend bekannte Angriffsmethoden beschleunigt, statt grundlegend neue zu ermöglichen.[19] Für Systeme wie MAKI, die auf einer Lieferkette aus Entwicklungskonsortium, externen Modellkomponenten und laufenden Updates beruhen (vgl. C.I.4.j), treten zusätzlich Risiken der Software-Lieferkette hinzu, etwa nicht nachvollziehbare Abhängigkeiten von Drittkomponenten. Davon betroffen sind alle drei Akteursebenen: auf der primären Ebene Verfahrensbeteiligte, auf der sekundären Ebene Justizinstitutionen und auf der tertiären Ebene das gesellschaftliche Vertrauen in die Integrität der Rechtsprechung. (Siehe zudem zu den Aspekten des Schutzes informationstechnischer Systeme vor internen und externen Störungen, einschließlich der Thematik der Prompt-Injections und der Sicherheit der Lieferkette die Abschnitte C.II.3.b.aa und C.II.3.b.dd sowie zur Cybersicherheit und KI-spezifischen Angriffen den Abschnitt C.II.3.b.cc)
Die vier beschriebenen Verschiebungen sind dabei nicht unabhängig voneinander. Neue Akteure, veränderte Interaktionsformen, systemische Kopplungen und technische Abhängigkeiten verstärken sich wechselseitig. Eine verantwortliche Risikoanalyse muss diesen Zusammenhang in den Blick nehmen – und damit die Grundlage für die Entwicklung von Evaluations- und Implementationskriterien legen, die im Folgenden entfaltet werden.

d. Opportunitätskosten der Nichtnutzung

Die bisherige Analyse hat Risiken primär als Folge der Entwicklung, Einführung und Nutzung von KI-Systemen betrachtet. Aus systemischer Sicht stellt sich jedoch

[17] *Anthropic*, Disrupting the first reported AI-orchestrated cyber espionage campaign, 13.11.2025, abrufbar unter https://www.anthropic.com/news/disrupting-AI-espionage (Abruf am 25.10.2026, 14:09 Uhr).
[18] *Google Threat Intelligence Group*, GTIG AI Threat Tracker: Advances in threat actor usage of AI tools, November 2025, abrufbar unter https://cloud.google.com/blog/topics/threat-intelligence/threat-actor-usage-of-ai-tools (Abruf am 25.10.2026, 14:11 Uhr).
[19] *OpenAI*, Disrupting malicious uses of AI: October 2025, 07.10.2025, abrufbar unter https://openai.com/global-affairs/disrupting-malicious-uses-of-ai-october-2025 (Abruf am 25.10.2026, 14:18 Uhr).

spiegelbildlich die Frage, ob auch die Nichtnutzung solcher Systeme als Risiko verstanden werden kann. Strukturelle Belastungen des Justizsystems durch Überlastung der Gerichte, lange Verfahrensdauern und begrenzte Ressourcen sind statistisch dokumentiert.[20] Wenn KI-Systeme unter bestimmten Bedingungen dazu beitragen könnten, Verfahren zu beschleunigen oder Ressourcen effizienter einzusetzen, wäre der Verzicht auf ihren Einsatz potentiell mit messbaren Kosten verbunden. In diesem Sinne wäre die Nichtnutzung ein eigenständiger Risikofaktor des Status quo.

Eine solche Argumentation setzt jedoch voraus, dass sich die behaupteten Effizienzgewinne empirisch belegen und zumindest grob quantifizieren lassen. Dafür wäre zunächst eine klare Bestimmung der relevanten Zielgrößen im Sinne von Leistungskennzahlen (Key Performance Indicators) erforderlich: etwa durchschnittliche Verfahrensdauer, Rückstände an unerledigten Verfahren, Belastung einzelner Spruchkörper oder Fehlerquoten in ausgewählten Verfahrensarten. Darüber hinaus müsste präzise definiert werden, an welcher Stelle im Prozess ein KI-System eingesetzt wird und welche Funktion es erfüllt, zum Beispiel in der Aktenzuordnung, in der Vorbereitung von Entscheidungsentwürfen oder in der Priorisierung von Fällen. Erst auf dieser Grundlage wäre eine vergleichende Analyse möglich, die zeigt, ob der Einsatz eines konkreten Systems bei vergleichbaren Rahmenbedingungen zu messbaren Verbesserungen führt. Die bloße Möglichkeit, dass KI-Systeme helfen könnten, genügt nicht, um von echten Opportunitätskosten zu sprechen.

Selbst wenn Effizienzgewinne empirisch nachweisbar wären, folgt daraus nicht automatisch, dass sie in zusätzliche Rechtsschutzleistung umgesetzt werden. Bekannte Phänomene wie die Ausdehnung von Tätigkeiten auf verfügbare Zeit (Parkinsonsches Gesetz) und Trittbrettfahrerverhalten legen nahe, dass frei werdende Kapazitäten auch in andere Aktivitäten fließen können, etwa in interne Verwaltungsaufgaben, informelle Qualitätssteigerung oder schlicht in Entlastung. Ob dies aus normativer Sicht problematisch ist, hängt wiederum von der Zielbestimmung richterlicher Tätigkeit ab. Eine rein quantitativ verstandene Leistungsmessung, die ausschließlich Fallzahlen als Indikator betrachtet, würde qualitative Gewinne wie z. B. sorgfältigere Begründungen, intensivere Beweiswürdigung, bessere Kommunikation mit Verfahrensbeteiligten systematisch ausblenden. Eine solche Messung wäre nicht nur verkürzt, sondern birgt ihrerseits Risiken wie eine problematische Steuerung richterlicher Tätigkeit über Kennzahlen, die den Anforderungen des Einzelfalls nicht gerecht wird.

Selbst wenn sich empirisch belegbare Opportunitätskosten der Nichtnutzung feststellen ließen, stellt sich die Frage nach den normativen Konsequenzen. Ein daraus abgeleiteter „Nutzungszwang“ für Gerichte oder einzelne Richterinnen und Richter wäre mit den bestehenden Vorstellungen richterlicher Unabhängigkeit nur schwer vereinbar und würde eine grundsätzliche Diskussion bisheriger Rechtsprechung zur richterlichen Organisations- und Technikfreiheit auslösen.[21] Wahrscheinlicher und verfassungsrechtlich möglicherweise weniger problematisch, aber nicht unbedenklich, wären weichere Steuerungsformen, etwa institutionelle Anreize, Priorisierung in der Ressourcenzuteilung oder die Vorgabe, bestimmte Werkzeuge „regelmäßig in Betracht zu ziehen“ (zur Thematik der Sicherstellung der

[20] Z. B. *Statistisches Bundesamt (Destatis)*, Statistischer Bericht – Zivilgerichte – 2025 (EVAS-Nummer 24231), 2026, abrufbar unter https://www.destatis.de/DE/Themen/Staat/Justiz-Rechtspflege/Publikationen/Downloads-Gerichte/statistischer-bericht-zivilgerichte-2100210257005.html (Abruf am 25.10.2026, 14:36 Uhr).

[21] Vgl. hierzu bereits *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 192ff.

Freiwilligkeit der Systemnutzung durch Richterinnen und Richter aus rechtlicher Perspektive siehe C.II.9). Solche Instrumente müssten jedoch sorgfältig austariert werden. Faktische Standardisierung von Entscheidungsverhalten über technische Systeme erzeugt ihrerseits neue systemische Risiken der Art, wie sie in den vorangegangenen Abschnitten beschrieben wurden.
Schließlich darf die Fokussierung auf Opportunitätskosten der Nichtnutzung nicht den Blick darauf verstellen, dass jeder Effizienzgewinn durch KI mit neuen Risiken erkauft werden kann. Zeitgewinne können mit erhöhten epistemischen Unsicherheiten, neuen Abhängigkeiten von technischen Infrastrukturen oder einer schleichenden Verschiebung der Verantwortungslasten einhergehen. Die Entscheidung, ein System nicht einzusetzen, kann daher im Einzelfall eine rationale Reaktion auf schwer quantifizierbare systemische Risiken sein. Der Begriff des „Risikos der Nichtnutzung“ ist nur dann analytisch tragfähig, wenn er symmetrisch zu den Risiken der Nutzung betrachtet wird und nicht als einseitiges Argument zugunsten technischer Aufrüstung fungiert.
Eine verantwortliche Justizpolitik muss beide Seiten berücksichtigen und offenlegen, auf welcher empirischen und normativen Grundlage sie Opportunitätskosten des Status quo gewichtet. Empirische Befunde zur Leistungsfähigkeit von KI-Systemen sind somit notwendig, aber nicht hinreichend. Ihre normative Bewertung bleibt eine eigenständige, explizit zu führende Aufgabe.

3. Methoden der Risikoanalyse

Die vorangegangenen Abschnitte haben gezeigt, dass Risiken beim KI-Einsatz in der Justiz unterschiedliche Formen annehmen: Sie können idiosynkratisch oder systemisch sein, sind probabilistisch schätzbar oder fundamental unbestimmt, und sie betreffen Akteursgruppen mit divergierenden Wahrnehmungen und Interessen. Eine Methodik der Risikoanalyse muss dieser Heterogenität gerecht werden. Nicht alle relevanten Risiken lassen sich ex ante identifizieren oder quantifizieren. Eine Methodik, die ausschließlich auf Quantifizierbarkeit setzt, würde gerade jene Risiken systematisch ausblenden, die am schwersten zu erfassen, aber potentiell am folgenreichsten sind.
Analytisch sinnvoll ist daher eine Kombination qualitativer und quantitativer Methoden, die unterschiedliche Funktionen erfüllen. Qualitative Methoden dienen primär der Rekonstruktion möglicher Wirkpfade und der Offenlegung zentraler Annahmen. Quantitative Methoden können helfen, Unsicherheiten explizit zu modellieren, Einflussfaktoren vergleichbar zu machen und potenzielle Interventionspunkte zu identifizieren. In der vorliegenden Analyse steht die qualitative Dimension im Vordergrund, die quantitative Analyse wird kurz eingeführt, aber nicht vertieft.
Den konzeptionellen Ausgangspunkt bildet die bereits eingeführte Unterscheidung verschiedener Formen von Unsicherheit, die sich für die Zwecke der Risikoanalyse in drei operative Kategorien übersetzen lässt. Die Grenzen der ex ante Risikoanalyse werden insbesondere in der dritten Kategorie sichtbar.

1. Probabilistisch schätzbare Risiken, etwa messbare Fehlerraten oder Ausfälle in klar abgegrenzten technischen Aufgaben, bei denen Eintrittswahrscheinlichkeiten prinzipiell bestimmbar sind.
2. Plausibilitätsbasierte Risiken, bei denen quantitative Schätzungen bspw. aufgrund fehlender Daten nicht möglich sind, qualitative Einschätzungen über Wirkzusammenhänge jedoch auf theoretischen Überlegungen, Erfahrungswissen oder analogen Kontexten beruhen können.
3. Fundamental unsichere Risiken, die emergent, systemisch oder stark kontextabhängig sind und sich weder sinnvoll quantifizieren noch zuverlässig qualitativ bewerten lassen.

Ein Teil relevanter Risiken bleibt dennoch ex ante unsichtbar. In komplexen sozio-technischen Systemen entstehen Risiken häufig erst aus der Interaktion technischer Komponenten, organisatorischer Routinen und menschlicher Entscheidungen. Solche Effekte sind retrospektiv häufig erkennbar, prospektiv jedoch schwer vorhersehbar. Vor diesem Hintergrund rückt die Robustheit institutioneller Systeme stärker in den Mittelpunkt der Analyse. Justizsysteme operieren nicht isoliert, sondern innerhalb organisatorischer, technischer und sozialer Netzwerke. Lokale Störungen können sich über solche Verflechtungen verstärken. Die zentrale Herausforderung besteht daher weniger in der exakten Vorhersage einzelner Risiken als in der Fähigkeit des Systems, unerwartete Belastungen zu absorbieren, ohne seine grundlegenden Funktionen zu verlieren.
Gleichwohl bleibt ein strukturiertes Risikomapping notwendig. Resilienzorientierung ersetzt keine systematische Analyse. Ohne sie lassen sich weder potenzielle Verwundbarkeiten identifizieren noch institutionelle Vorsorgemechanismen entwickeln. Risikomodelle sind jedoch heuristische Konstruktionen, keine Abbildungen der Wirklichkeit. Ihre Aussagekraft hängt entscheidend davon ab, ob die zugrunde liegenden Annahmen, Unsicherheiten und Leerstellen transparent gemacht werden. Dies erfordert einen methodisch reflektierten Umgang mit empirischen Befunden und

eine explizite Auseinandersetzung mit Reichweite und Grenzen der verwendeten Risikomodelle.
Gerade in komplexen institutionellen Kontexten erfordert dies einen interdisziplinären Zugang. Richterliche Entscheidungsfindung verbindet rechtliche Normstrukturen, organisationale Prozesse, menschliche Urteilsbildung und technische Infrastrukturen. Eine Risikoanalyse muss diese Ebenen gemeinsam betrachten. Risikomapping ist daher kein einmaliges Ergebnis, sondern ein iterativer Prozess. Modelle dienen als vorläufige Orientierungen, die mit neuen empirischen Erkenntnissen revidiert und verfeinert oder sogar korrigiert werden müssen.

a. Qualitative Risikoanalyse

Die qualitative Risikoanalyse konzentriert sich nicht auf die Messung von Eintrittswahrscheinlichkeiten, sondern auf die Modellierung und Rekonstruktion möglicher Wirkpfade und die Offenlegung von strukturellen Abhängigkeiten.
Systemische Risiken lassen sich nicht allein über beobachtbare Fehlereignisse erfassen. Sie resultieren nicht primär aus einzelnen Fehlern, sondern entstehen aus der Struktur und Vernetzung eines Systems. Lokale Störungen können sich über Interdependenzen und Rückkopplungen verstärken und Wirkungen entfalten, die in keinem proportionalen Verhältnis zu ihrem ursprünglichen Auslöser stehen.[22] Solche Dynamiken sind aus der Analyse komplexer Netzwerke bekannt. Forschung zu Finanz- und Produktionssystemen zeigt, dass steigende Komplexität häufig mit Zentralisierung einhergeht. Diese kann Effizienzgewinne erzeugen, zugleich jedoch neue Verwundbarkeiten schaffen, weil Störungen an zentralen Knoten systemweit propagieren können.[23] Das „robust-yet-fragile“-Muster *legt eine methodische Konsequenz nahe*: Qualitative Risikoanalyse nimmt genau jene Strukturen in den Blick, die unter Normalbedingungen unsichtbar bleiben, weil sie nur unter Belastung zum Risikofaktor werden und sich statistischen Methoden daher entziehen.
Auch das Justizsystem weist entsprechende Eigenschaften auf. Richterliche Entscheidungen entstehen nicht isoliert, sondern innerhalb eines Zusammenspiels individueller Entscheidungspraktiken, organisatorischer Routinen und technischer Infrastrukturen. Der Einsatz von KI-Systemen kann diese Struktur in mehrerlei Hinsicht verändern. Zum einen können Modellannahmen, Trainingsdaten oder algorithmische Heuristiken die Entscheidungsgrundlagen in großer Breite prägen, wenn viele Gerichte auf identische oder ähnliche Systeme zurückgreifen. Fehlerhafte Annahmen oder systematische Verzerrungen reproduzieren sich dann über zahlreiche Verfahren hinweg. Darüber hinaus entstehen infrastrukturelle Abhängigkeiten. Der Betrieb von KI-Systemen erfordert spezialisierte Software, Dateninfrastrukturen und Wartung, die häufig von einzelnen Anbietern bereitgestellt werden. Gerichte können dadurch von bestimmten Systemarchitekturen, Updatezyklen oder Datenformaten abhängig werden. Änderungen am System können dann unmittelbare Auswirkungen auf Entscheidungsprozesse haben, ohne vollständig transparent zu sein. Letztlich können sich institutionelle Routinen dahingehend verschieben, dass algorithmische Empfehlungen faktisch zum Referenzpunkt für Entscheidungen werden, sodass

[22] *Elliott/Golub*, Networks and Economic Fragility, Annual Review of Economics 14 (2022), 665; *Piccardi/Tajoli*, Complexity, centralization, and fragility in economic networks, PLOS ONE 13 (2018), e0208265.
[23] *Tonzer*, Cross-border interbank networks, banking risk and contagion, Journal of Financial Stability 18 (2015), 19; *Piccardi/Tajoli*, Complexity, centralization, and fragility in economic networks, in: PLOS ONE 13 (2018), e0208265.

Abweichungen erklärungsbedürftig sind, obwohl die formale Entscheidungsverantwortung weiterhin beim Gericht liegt.
Auch hier wird deutlich, dass Wirkungen nicht aus einzelnen Ursachen entstehen, sondern aus dem Zusammenwirken mehrerer Faktoren, deren Interaktionen das tatsächliche Risikoprofil bestimmen. Eine qualitative Risikoanalyse muss daher die Wirkungsannahmen explizieren, die solchen Implementationen zugrunde liegen. Jede Analyse setzt implizit voraus, dass bestimmte Handlungen oder Technologien bestimmte Effekte erzeugen. Werden diese Annahmen nicht offengelegt, besteht die Gefahr systematischer Fehlzuschreibungen und einer Überschätzung kausaler Zusammenhänge.[24] Besonders problematisch ist die verbreitete Praxis, komplexe Ereignisse vorschnell auf einen einzelnen plausiblen Ursache-Wirkungs-Zusammenhang zurückzuführen. In der Praxis werden statistische Korrelationen oder qualitative Plausibilitäten häufig als hinreichende Erklärung interpretiert, obwohl in komplexen Entscheidungssystemen meist mehrere Faktoren gleichzeitig wirken. Arbeiten aus der Risiko- und Entscheidungsforschung zeigen, dass solche Vereinfachungen anfällig für Expertenbias, kognitive Verzerrungen und motiviertes Schließen sind.[25] Gerade in institutionellen Kontexten entstehen Risiken häufig nicht durch isolierte Ursachen, sondern durch das Zusammenwirken mehrerer Faktoren, deren Interaktionen entscheidend für das tatsächliche Risikoprofil sein können.
Das Ziel der qualitativen Risikoanalyse ist daher, mögliche Wirkpfade zu rekonstruieren, ohne deren Eintritt oder Stärke vorhersagen zu wollen. Entscheidend ist, strukturelle Abhängigkeiten sichtbar zu machen, die in rein ereignisorientierten Betrachtungen verborgen bleiben. Ein Beispiel verdeutlicht diese Logik. Wird ein KI-System eingeführt, das Empfehlungen zur Strafzumessung generiert, lassen sich unterschiedliche Wirkannahmen formulieren: Die Empfehlung könnte als kognitiver Anker wirken und die Streuung von Strafhöhen reduzieren; sie könnte als bewusste Orientierungshilfe genutzt werden, ohne die richterliche Bewertung zu ersetzen; oder organisatorische Effizienzerwartungen könnten ihre faktische Bindungswirkung erhöhen, ohne dass dies formell vorgesehen ist. Diese unterschiedlichen Annahmen führen zu verschiedenen Risikobewertungen. Entscheidend ist daher nicht ihre unmittelbare empirische Bestätigung, sondern ihre transparente Formulierung und Kommunikation zwischen den beteiligten Akteuren.
Entwickler, Justizverwaltung und richterliche Praxis müssen offenlegen, welche Wirkungen sie erwarten, welche Unsicherheiten bestehen und welche institutionellen Schutzmechanismen vorgesehen sind. Erst eine solche explizite Reflexion von Wirkannahmen ermöglicht eine überprüfbare und lernfähige Risikosteuerung.

b. Quantitative Risikoanalyse

Quantitative Risikoanalyse baut auf den in der qualitativen Analyse identifizierten Wirkannahmen auf. Ihr Ziel besteht nicht darin, komplexe institutionelle Systeme vollständig zu quantifizieren. Vielmehr dient sie dazu, Unsicherheit strukturiert zu modellieren, Einflussfaktoren vergleichbar zu machen und die Konsequenzen unterschiedlicher Annahmen sichtbar zu machen. Beide Ansätze sind damit komplementär.
Quantitative Modelle können dabei Unsicherheit explizit repräsentieren, statt sie implizit zu belassen, und ermöglichen die vergleichende Analyse möglicher

[24] *Cox*, Improving Causal Inferences in Risk Analysis, Risk Analysis 33 (2013), 1762.
[25] *Cox*, Improving Causal Inferences in Risk Analysis, Risk Analysis 33 (2013), 1762; *Cox/Ricci*, Causation in risk assessment and management: models, inference, biases, and a microbial risk-benefit case study, Environment International 31 (2005), 377.

Einflussfaktoren und damit die Identifikation plausibler Interventionspunkte. Ferner machen sie Annahmen überprüfbar, indem sichtbar wird, welche Schlussfolgerungen von welchen Prämissen abhängen.[26] Für diese Zwecke kommen verschiedene Modellierungsansätze in Betracht. Welcher Ansatz geeignet ist, hängt von Datenlage, Zeithorizont und Risikotyp ab.

Im Kontext der richterlichen Entscheidungsfindung sind folgende fünf Ansätze plausibel. **Bayesianische Netzwerke** modellieren bedingte Abhängigkeiten zwischen Variablen und erlauben probabilistische Schlussfolgerungen unter Unsicherheit. Sie sind besonders geeignet, wenn Risiken aus verketteten technischen, organisatorischen und menschlichen Faktoren entstehen und wenn empirische Daten mit Expertenschätzungen kombiniert werden müssen.[27] **Szenario- und Stressmodelle** verzichten zumeist auf eine vollständige Wahrscheinlichkeitsparametrisierung und konzentrieren sich stattdessen auf das Verhalten eines Systems unter definierten Belastungen. Sie sind besonders nützlich in Situationen fundamentaler Unsicherheit, in denen es primär um Robustheit, Eskalationspfade und institutionelle Verwundbarkeiten geht. **Fehlerbaum- und Ereignisbaumanalysen** strukturieren mögliche Ketten von Ausfällen und Sicherheitsbarrieren. Sie eignen sich insbesondere für technische oder prozessuale Risiken, etwa Ausfälle digitaler Infrastruktur, fehlerhafte Datenübertragung oder Störungen einzelner Systemkomponenten. **Reliabilitäts- und Verfügbarkeitsmodelle** quantifizieren Ausfallraten, Wiederherstellungszeiten und infrastrukturelle Abhängigkeiten. Solche Modelle adressieren Risiken der Betriebsfähigkeit, die im Justizkontext relevant werden, wenn Verfahren von digitalen Plattformen oder zentralen Dateninfrastrukturen abhängig sind. **Statistische Outcome-Modelle** schließlich analysieren beobachtbare Ergebnisverteilungen, etwa Fehlerraten, Verschiebungen von Entscheidungsmustern oder ungleiche Effekte auf verschiedene Gruppen. Diese Modelle sind besonders geeignet, wenn der Fokus auf empirisch messbaren Outputs liegt. Sie erfassen jedoch nur begrenzt die institutionellen Mechanismen, die zu diesen Ergebnissen führen.

Gemeinsam ist diesen Ansätzen, dass sie Unsicherheit nicht eliminieren, sondern strukturieren. Ihr Wert liegt darin, plausible Risikohebel sichtbar zu machen und Monitoring sowie Governance auf jene Punkte zu konzentrieren, an denen Entscheidungen das Risikoprofil tatsächlich beeinflussen.

Zwei Beispiele verdeutlichen diese Logik: Eine Szenarioanalyse könnte untersuchen, ob KI-Empfehlungen zur Strafzumessung langfristig zu einer Standardisierung richterlicher Entscheidungen führen, indem verschiedene Nutzungsintensitäten modelliert werden, von einer losen Orientierungshilfe bis hin zu einer faktischen Standardisierung. Für jedes Szenario lassen sich Parameter definieren, etwa der Anteil der Verfahren mit KI-Empfehlung, die Wahrscheinlichkeit der Übernahme einer Empfehlung oder die durchschnittliche Abweichungsrate. Ziel ist dabei nicht die exakte Prognose, sondern die Analyse der Bedingungen, unter denen eine lokale Entscheidungshilfe zu systemweiten Effekten führen könnte. Eine Fehlerbaumanalyse könnte demgegenüber untersuchen, unter welchen Kombinationen von Ereignissen ein Systemausfall den Verfahrensbetrieb stört. Mögliche Ereignisse sind etwa ein Ausfall der Serverinfrastruktur, fehlerhafte Datenübertragung, fehlende lokale Fallback-Prozesse oder unzureichende Schulung der Mitarbeitenden im Umgang mit Systemstörungen. Auf dieser Grundlage kann analysiert werden, welche

[26] Vgl. allgemein *Fenton/Neil*, Risk Assessment and Decision Analysis with Bayesian Networks, 1. Aufl. 2012.

[27] Vgl. allgemein *Korb/Nicholson*, Bayesian Artificial Intelligence, 2. Aufl. 2010; *Fenton/Neil*, Risk Assessment and Decision Analysis with Bayesian Networks, 1. Aufl. 2012.

Schutzmaßnahmen, etwa redundante Infrastruktur, Offline-Fallbacks oder organisatorische Notfallprotokolle, das Gesamtrisiko am effektivsten reduzieren.
Im Evaluationsrahmen (Kapitel C.I.4) finden diese Ansätze vor allem im laufenden Betrieb ihren Platz: Die vorgesehenen Monitoring-Daten sind die Grundlage für statistische Outcome-Modelle und Szenarioanalysen zur Standardisierung, während Fehlerbaum- und Verfügbarkeitsanalysen die infrastrukturellen Risiken adressieren, die der Rahmen selbst nur selektiv berücksichtigt.

c. Grenzen quantitativer Risikoanalyse

Quantitative Risikoanalyse erweitert das analytische Instrumentarium, hebt jedoch die grundlegenden epistemischen Grenzen komplexer sozio-technischer Systeme nicht auf. Ohne methodische Vorsicht können quantitative Modelle selbst neue Fehlinterpretationen erzeugen und damit jene Transparenz untergraben, die sie eigentlich herstellen sollen.
Die Grenzen sind vielfältig. Erstens besteht eine strukturelle Abhängigkeit von Modellannahmen: Jedes quantitative Modell setzt voraus, dass bestimmte Einflussfaktoren ausgewählt und bestimmte Beziehungen zwischen ihnen angenommen werden. Werden relevante Faktoren ausgelassen oder falsch modelliert, können auch formal korrekte Berechnungen systematisch irreführend sein.[28] Dieses Problem ist besonders ausgeprägt in institutionellen Kontexten, in denen organisatorische Routinen, informelle Praktiken oder kulturelle Erwartungen schwer messbar sind. Hinzu kommt die Gefahr der Scheingenauigkeit. Numerische Ergebnisse können den Eindruck erzeugen, Risiken seien präzise messbar. Tatsächlich beruhen viele Parameter auf begrenzten Daten oder Expertenschätzungen. Besonders bei seltenen, aber gravierenden Risiken fehlen stabile Referenzklassen. Quantitative Modelle sollten daher primär als Darstellung von Unsicherheitsräumen verstanden werden, nicht als Instrument exakter Vorhersage.[29]
Ein weiteres Problem betrifft Datenknappheit und die Rolle von Expertenschätzungen. In vielen Anwendungen übersteigt die Zahl potenzieller Einflussfaktoren die verfügbare Datenbasis. Modelle greifen deshalb häufig auf Expertenschätzungen zurück. Diese sind methodisch zulässig, erhöhen jedoch die Anfälligkeit für kognitive Verzerrungen, institutionelle Interessen oder Ankereffekte. Transparente Dokumentation der Herkunft und Unsicherheit solcher Parameter ist daher unverzichtbar. Auch Sensitivitätsanalysen haben Grenzen. Sie können zeigen, welche Variablen innerhalb eines Modells einen großen Einfluss auf Ergebnisse haben. Dieser Einfluss ist jedoch relativ zur gewählten Modellstruktur. Ein stark sensitiv wirkender Parameter muss daher nicht zwingend der wichtigste Faktor im realen System sein. Hinzu kommt die Zeitdimension. Viele Risiken entstehen nicht punktuell, sondern kumulativ über längere Zeiträume, etwa durch institutionelles Lernen, Standardisierung oder veränderte Erwartungshaltungen. Statische Modelle können solche Dynamiken nur begrenzt erfassen.
Hinzu kommt eine normative Grenze, die sich von den beschriebenen Grenzen unterscheidet. Quantitative Analysen können Entscheidungen strukturieren, aber diese nicht treffen. Die Auswahl von Zielgrößen, Schwellenwerten oder akzeptablen Risikoniveaus ist keine technische, sondern eine normative Entscheidung. Wird dieser Unterschied verwischt, besteht die Gefahr einer Technokratisierung normativer

[28] *Cox*, Improving Causal Inferences in Risk Analysis, Risk Analysis 33 (2013), 1762.
[29] Vgl. allgemein *Fenton/Neil*, Risk Assessment and Decision Analysis with Bayesian Networks, 1. Aufl. 2012.

Fragen, bei der Zahlen politische oder rechtliche Wertungen verdecken, statt sie transparent zu machen.
Quantitative Risikoanalyse ist daher im vorliegenden Kontext als ergänzendes Instrument zu verstehen. Ihre Stärke liegt in der strukturierten Exploration von Abhängigkeiten und Szenarien. Ihre Grenzen liegen dort, wo formale Präzision epistemische Unsicherheit überdeckt oder normative Entscheidungen als rein technische Probleme erscheinen lässt.

4. Evaluations- und Reflexionsrahmen für den KI-Einsatz

a. Zweck des Evaluationsrahmens

Die vorangegangenen Kapitel haben gezeigt, dass Risiken beim Einsatz von KI-Systemen in der richterlichen Entscheidungsfindung nicht allein aus technischen Eigenschaften eines Systems entstehen. Sie ergeben sich vielmehr aus dem Zusammenspiel von Entscheidungsfunktion, institutionellem Kontext, organisatorischer Einbettung und der Art, wie menschliche Entscheidungsträger algorithmische Informationen nutzen und interpretieren. Eine belastbare Bewertung des KI-Einsatzes erfordert daher ein Instrument, das diese Faktoren systematisch sichtbar und ihre Bedeutung für konkrete Anwendungssituationen prüfbar macht.
Der folgende Evaluations- und Reflexionsrahmen verfolgt dieses Ziel. Er überführt die analytischen Perspektiven der vorangegangenen Kapitel – die Risikotypologie, die Stakeholderdifferenzierung, die methodischen Grundsätze der Risikoanalyse – in Reflexionsfragen, mit deren Hilfe der Einsatzkontext eines KI-Systems analysiert werden kann. Auf dieser Grundlage lassen sich anschließend kontextspezifische Risikoszenarien entwickeln und geeignete Mitigationsmaßnahmen identifizieren.
Der Rahmen konzentriert sich dabei auf Risiken, die aus der Interaktion zwischen Anwendern und KI-Systemen im konkreten Entscheidungsprozess hervorgehen und sich unmittelbar auf der primären Akteursebene auswirken, also bei den Verfahrensbeteiligten und den entscheidenden Richterinnen und Richtern. Andere Risikoperspektiven, etwa infrastrukturelle Abhängigkeiten oder großskalige Systemvernetzungen, werden nur selektiv berücksichtigt. Im Mittelpunkt stehen Risiken, die aus der Struktur der Entscheidungssituation selbst hervorgehen, insbesondere aus der Art der zu treffenden Entscheidung, der Rolle des Systems im Entscheidungsprozess sowie den Handlungsmöglichkeiten der beteiligten Akteure.
Diese Fokussierung ist analytisch begründet, nicht normativ. Sie bedeutet nicht, dass infrastrukturelle oder systemische Risiken nachrangig wären – die vorangegangenen Kapitel haben deren Bedeutung ausführlich entwickelt. Sie bedeutet, dass der Evaluationsrahmen dort ansetzt, wo Risiken für die primär Betroffenen am unmittelbarsten wirksam werden, und von dort aus zu weiteren Ebenen ausgreift. Die daraus entwickelten Risikoszenarien sind daher als Funktion der jeweiligen Entscheidungsumwelt zu verstehen: Der Kontext, etwa der Grad normativer Abwägung, die Komplexität der Entscheidung oder die Einbindung algorithmischer Empfehlungen, bestimmt maßgeblich, welche Risiken plausibel sind und welche Formen der Risikominderung sinnvoll erscheinen.

b. Risikodimensionen des KI-Einsatzes

Bevor kontextspezifische Risikoszenarien entwickelt werden können, ist eine Einordnung der zu bewertenden Anwendung entlang zentraler Risikodimensionen erforderlich. Diese Dimensionen bestimmen, welche Risiken im konkreten Fall überhaupt plausibel sind. Sie bilden die analytische Vorstufe zu jeder weiterführenden Bewertung. Aufbauend auf den im ersten Bericht dieses Projekts entwickelten Überlegungen[30] lassen sich drei solche Dimensionen unterscheiden: der funktionale Charakter der Tätigkeit, der Grad der Automatisierung sowie die Komplexität des Entscheidungsfeldes.

[30] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025.

Funktionaler Charakter der Tätigkeit
Die erste Dimension betrifft die Art der Tätigkeit, auf die sich der KI-Einsatz bezieht. Administrative Funktionen, etwa Dokumentenstrukturierung oder Informationsaufbereitung, greifen typischerweise nicht unmittelbar in rechtliche Bewertungen ein und sind daher in der Regel mit geringeren Risiken verbunden, auch wenn sie Verfahrensrechte berühren können (vgl. C.I.4.c.aa). Inhaltlich-normative Tätigkeiten betreffen dagegen zentrale Elemente richterlicher Entscheidungsfindung, etwa die Bewertung von Sachverhalten, rechtliche Subsumtion oder Entscheidungsbegründungen. In diesen Bereichen steigt die ethische und institutionelle Relevanz des KI-Einsatzes deutlich, weil Fehler oder Verzerrungen des Systems unmittelbar in die Entscheidungssubstanz eingreifen.

Grad der Automatisierung
Die zweite Dimension beschreibt die Rolle des Systems im Entscheidungsprozess. KI-Systeme können Informationen bereitstellen, Entscheidungsprozesse strukturieren oder konkrete Empfehlungen generieren. Je stärker ein System Vorentscheidungen trifft oder den Ausgang eines Entscheidungsprozesses faktisch vorstrukturiert, desto größer wird das Risiko einer Entscheidungsvorverlagerung auf das System. Dabei ist der Grad der Automatisierung nicht allein technisch zu bestimmen, sondern auch organisatorisch, da institutionelle Arbeitsbedingungen die tatsächliche Nutzung eines Systems maßgeblich beeinflussen. Der Bericht verwendet ‚Automatisierungsgrad‘ als Oberbegriff; ‚Delegationsgrad‘ bezeichnet dessen technische Seite, also die Rolle, die ein System nach seiner Gestaltung übernimmt (C.I.4.f.cc), ‚Vorstrukturierung‘ die faktische Wirkung auf den Entscheidungsprozess unabhängig von dieser Rolle.
Diese graduelle Einordnung lässt sich auch mit einer binären Unterscheidung zwischen entscheidungsersetzendem und entscheidungsunterstützendem Einsatz verbinden: Ein entscheidungsersetzender Einsatz entspricht dem oberen Ende dieser Skala, bei dem das System selbst eine letztverbindliche Entscheidung trifft; ein entscheidungsunterstützender Einsatz umfasst demgegenüber das gesamte Spektrum von informierenden über strukturierende bis hin zu empfehlenden Funktionen, bei denen die Letztentscheidung beim Richter oder der Richterin verbleibt.

Komplexität des Entscheidungsfeldes
Die dritte Dimension betrifft die Struktur des jeweiligen Anwendungsbereichs. In stark formalisierten und wiederkehrenden Entscheidungssituationen mit klar definierten Kriterien kann technische Unterstützung vergleichsweise gut integriert werden. Demgegenüber stehen hochkomplexe Entscheidungskontexte, die durch offene Bewertungsmaßstäbe und einzelfall- sowie kontextabhängige Abwägungen geprägt sind. In solchen Situationen besteht ein erhöhtes Risiko, dass algorithmische Systeme durch Vereinfachung oder implizite Modellannahmen unbeabsichtigte Verzerrungen erzeugen.

Das Risikopotenzial eines KI-Systems ergibt sich aus der Kombination dieser drei Dimensionen. Besonders kritische Konstellationen entstehen dort, wo eine inhaltlich-normative Tätigkeit, ein hoher Grad der Automatisierung und ein komplexes Entscheidungsfeld zusammentreffen.
Der Rahmen wird in drei Phasen angewendet: Kontexteinordnung aus Anwenderperspektive (C.I.4.c), technische Gestaltung aus Designerperspektive (C.I.4.f) und institutionelle Einbettung aus organisationaler Perspektive (C.I.4.g). Die erste Phase des Evaluationsrahmens dient der Kontextanalyse. Ihr Ziel ist die Einordnung einer Anwendung entlang dieser drei Dimensionen. Erst wenn diese

Grundkonstellation geklärt ist, lässt sich beurteilen, welche Risiken im konkreten Fall plausibel sind und welche Form weiterer Analyse erforderlich ist. Die nachfolgenden Anwenderfragen operationalisieren diese Dimensionen auf der Ebene der konkreten Entscheidungssituation. Sie erfassen insbesondere, welche Güter betroffen sind, wie eingriffsintensiv und reversibel Entscheidungsfolgen sind, welche Fähigkeiten eine Tätigkeit voraussetzt und in welchem Umfang sie sich formalisieren oder algorithmisch unterstützen lässt. Auf dieser Grundlage bauen anschließend die Perspektiven der Systemdesigner und Organisationen auf, die sich mit der technischen Ausgestaltung und institutionellen Einbettung des bereits eingeordneten KI-Einsatzes befassen. Für jede identifizierte Risikoklasse des KI-Einsatzes ist das korrespondierende Risikoprofil des Status quo zu benennen. Wenn keine empirische Evidenz über den Status quo vorliegt, ist dies als Unvollständigkeit der Analyse auszuweisen.

c. Reflexions- und Evaluationsfragen für Anwender

Die Anwenderperspektive betrifft diejenigen Akteure, die Entscheidungen unter Nutzung eines KI-Systems treffen oder vorbereiten. Im Justizkontext sind dies insbesondere Richterinnen und Richter, Rechtspflegerinnen und Rechtspfleger sowie weitere Verfahrensakteure, die mit algorithmisch erzeugten Informationen oder Empfehlungen arbeiten.
Die folgenden Fragen klären, in welchem Kontext ein KI-System eingesetzt werden soll und ob dieser Einsatz mit den Anforderungen der jeweiligen Entscheidungsfunktion vereinbar ist. Sie folgen der Struktur der drei Risikodimensionen: dem funktionalen Charakter der Tätigkeit, dem Grad der Automatisierung sowie der Komplexität des Entscheidungsfeldes.

aa. Funktionaler Charakter der Tätigkeit

Die in diesem Abschnitt formulierten Fragen zielen primär auf die Risikodimension des funktionalen Charakters der Tätigkeit. Bewertet wird dabei die Tätigkeit, die das System unterstützt; ‚Entscheidung‘ bezeichnet in den folgenden Fragen stets die gerichtliche Entscheidung, auf die sich diese Tätigkeit auswirkt, nicht einen Output des Systems. Ob eine Tätigkeit administrativ oder inhaltlich-normativ ist, bestimmt sich – entsprechend der Unterscheidung im ersten Arbeitspaket – danach, ob sie den normativen Kern der richterlichen Entscheidungsfindung berührt, also mit der rechtlichen Bewertung, Argumentation oder Entscheidungsvorbereitung verbunden ist, oder ob sie lediglich der Organisation, Strukturierung oder technischen Unterstützung gerichtlicher Abläufe dient.[31] Die folgenden Fragen ersetzen diese Abgrenzung nicht. Sie dienen dazu, den Grad der Normativität der betroffenen Entscheidung und damit das Gewicht dieser Einordnung zu bestimmen: Je stärker eine Entscheidung subjektive Rechte, Grundfreiheiten, geschützte Interessen, Teilhabechancen oder die körperliche beziehungsweise psychische Integrität berührt, desto höher ist das Risiko, wenn ein System auf ihren normativen Kern einwirkt, und desto sorgfältiger ist zu prüfen, ob die unterstützte Tätigkeit diesen Kern tatsächlich unberührt lässt. Gleiches gilt für Entscheidungen mit hoher Eingriffsintensität, besonderer rechtlicher Bindung, asymmetrischen Machtverhältnissen oder personengebundener Verantwortung.
Zugleich liefern die Fragen erste Hinweise auf die beiden anderen Risikodimensionen. Entscheidungen mit hoher Eingriffsintensität und starkem Einzelfallbezug sind häufig nicht nur normativ sensibel, sondern auch komplex, da sie kontextsensitives Abwägen

[31] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 73f.

und individuelle Begründungen erfordern. Demgegenüber können abstrakt-generelle oder gruppenbezogene Entscheidungen zwar ebenfalls normativ bedeutsam sein, weisen jedoch teilweise stärker standardisierbare Strukturen auf.
Zu prüfen sind insbesondere die betroffenen Güter:

- Betrifft die Entscheidung subjektive Rechte, Grundfreiheiten oder rechtlich geschützte Interessen von Verfahrensbeteiligten?
- Hat sie Auswirkungen auf wirtschaftliche oder soziale Teilhabechancen, etwa durch finanzielle Belastungen oder berufliche Konsequenzen?
- Berührt sie die körperliche oder psychische Integrität, etwa durch Freiheitsentzug, medizinische Maßnahmen oder andere schwerwiegende Eingriffe?
- Kann eine fehlerhafte Systemfunktion Verfahrensrechte der Beteiligten berühren, etwa das rechtliche Gehör oder den Zugang zum Gericht, auch wenn sie die Sachfrage selbst nicht betrifft?

Eng damit verbunden ist die Frage nach der Eingriffsintensität der Entscheidung:

- Liegt eine Entscheidung mit erheblicher Eingriffsintensität vor?
- Ist die Entscheidung normativ stark gebunden, etwa durch gesetzliche Vorgaben oder verfassungsrechtliche Schutzprinzipien?

Entscheidend ist auch die Position der entscheidenden Person:

- Handelt es sich um eine Rolle, die rechtlich besonders reguliert oder mit spezifischen Verantwortlichkeiten verbunden ist?
- Besteht eine asymmetrische Machtposition zwischen Entscheidungsinstanz und Betroffenen?

Darüber hinaus ist zu klären, auf welcher Entscheidungsebene sich die Tätigkeit bewegt:

- Handelt es sich um eine individuell-konkrete Einzelfallentscheidung?
- Um eine gruppenbezogene Entscheidung, die mehrere Personen betrifft?
- Oder um eine abstrakt-generelle Steuerungsentscheidung, die institutionelle Prozesse oder Entscheidungsregeln prägt?

Die einzelnen Aspekte betreffen unterschiedliche rechtliche und ethische Dimensionen, die sich nicht sinnvoll addieren oder gegeneinander aufrechnen lassen. Ziel ist vielmehr eine begründete Einordnung des funktionalen Charakters der betreffenden Tätigkeit. Entscheidend ist dabei eine fachliche Gesamtbewertung der Entscheidungssituation. Besonderes Gewicht kommt Faktoren zu, die unmittelbar Grundrechte, erhebliche Eingriffsintensität oder ausgeprägte Machtasymmetrien betreffen. Dabei ist zu unterscheiden: Für die Einordnung als inhaltlich-normative Tätigkeit maßgeblich ist, ob die Tätigkeit die Beurteilung des geltend gemachten Rechts selbst berührt. Verfahrensgrundrechte, insbesondere der Anspruch auf rechtliches Gehör (Art. 103 Abs. 1 GG) und auf effektiven Rechtsschutz, können dagegen auch durch administrative Funktionen beeinträchtigt werden, etwa wenn ein System einen Schriftsatz fehlerhaft zuordnet und dieser deshalb unberücksichtigt bleibt. Eine solche Betroffenheit macht eine Tätigkeit nicht zu einer inhaltlich-normativen, ist aber bei der Bewertung ihres Risikos zu berücksichtigen. Wenn mehrere dieser Merkmale vorliegen, spricht dies regelmäßig für ein hohes Risiko bei der Einbindung technischer Unterstützung, insbesondere wenn die Tätigkeit inhaltlich-normativ ist. Die Fragen dienen somit als strukturierte Orientierung für eine besonnene

fachliche Einschätzung, die den normativen Charakter der Entscheidung transparent begründet.

bb. Reversibilität von Entscheidungsfolgen

Die Reversibilität von Entscheidungsfolgen wirkt innerhalb der Kontextanalyse als Verstärkungsfaktor. Je schwerer sich die Folgen einer Entscheidung nachträglich korrigieren lassen, desto stärker steigt das Risiko einer fehlerhaften technischen Vorstrukturierung und desto größer ist die Bedeutung einer sorgfältigen menschlichen Urteilsbildung vor der Entscheidung. Dabei genügt es nicht, auf formale Korrekturmöglichkeiten zu verweisen. Entscheidend ist, ob faktisch entstandene Schäden wie Zeitverluste, Reputationsbeeinträchtigungen, wirtschaftliche Belastungen oder Freiheitsbeschränkungen vollständig rückgängig gemacht werden können. Diese Frage ist für die spätere Beurteilung einer angemessenen Automatisierungstiefe zentral, weil irreversible Fehler das Gewicht jedes einzelnen Fehlentscheidungsrisikos erheblich erhöhen.
Zu prüfen ist daher:

- Sind die Folgen einer Entscheidung vollständig reversibel, etwa durch einfache Korrektur oder nachträgliche Anpassung?
- Sind sie nur teilweise reversibel, weil eine Korrektur mit zeitlichen Verzögerungen oder zusätzlichen Belastungen verbunden ist?
- Oder sind sie faktisch irreversibel, weil selbst eine spätere Korrektur die ursprünglichen Folgen nicht mehr vollständig aufheben kann?

Besondere Aufmerksamkeit verdienen Konstellationen, in denen eine spätere Korrektur zwar formal möglich ist, faktisch jedoch nicht alle Schäden beseitigt. Dies kann etwa der Fall sein bei erheblichem Zeitverlust, wirtschaftlichen Schäden, Reputationsverlusten oder bei Eingriffen wie Freiheitsentzug, Kündigung oder medizinischen Maßnahmen.
Die Antworten auf diese Fragen lassen sich entlang einer ordinalen Risikologik einordnen. Je schwerer die Folgen einer Entscheidung korrigiert werden können oder je höher die mit einer Korrektur verbundenen Kosten sind, desto höher ist das Risikopotenzial des KI-Einsatzes in diesem Kontext. Vollständig reversible Entscheidungen weisen ein deutlich geringeres Risiko auf als Konstellationen, in denen Korrekturen mit erheblichen zeitlichen, wirtschaftlichen oder sozialen Belastungen verbunden sind. Am oberen Ende der Skala stehen faktisch irreversible Entscheidungen, bei denen eine spätere Korrektur die ursprünglichen Folgen nicht mehr vollständig beseitigen kann. Die Reversibilität von Entscheidungsfolgen wirkt damit als gradueller Verstärkungsfaktor der Normativitätsdimension. Je irreversibler eine Entscheidung ist oder je höher die Reversibilitätskosten ausfallen, desto vorsichtiger ist eine technische Vorstrukturierung des Entscheidungsprozesses zu beurteilen.

cc. Erforderliche Fähigkeiten und Kenntnisse

Die Fragen dieses Abschnitts operationalisieren in erster Linie die Komplexität des Entscheidungsfeldes; einzelne Merkmale, insbesondere die personengebundene Verantwortung, berühren zugleich die Normativität der Tätigkeit. Sie prüfen, ob die betreffende Tätigkeit Fähigkeiten voraussetzt, die sich nur begrenzt in formalisierbare Entscheidungsregeln übersetzen lassen. Je stärker eine Entscheidung

kontextsensitives Abwägen, kommunikative Urteilskraft, Empathie oder personengebundene Verantwortung erfordert, desto eher handelt es sich um ein komplexes und normativ sensibles Entscheidungsfeld. In solchen Konstellationen spricht bereits die Struktur der Aufgabe gegen eine weitgehende technische Vorstrukturierung entscheidungsrelevanter Wertungen.
Der Abschnitt dient zugleich als Prüfung der Algorithmisierbarkeit. Wo rechtfertigende Begründung, situatives Urteil und personale Verantwortung zum Kern der Tätigkeit gehören, können technische Systeme in der Regel nur Teilaspekte unterstützen. Für die Kontextanalyse ist dies entscheidend, weil sich hier Aufgaben identifizieren lassen, bei denen die Natur der Tätigkeit selbst Grenzen sinnvoller Automatisierung setzt.
Zu prüfen ist:

- Setzt die betreffende Tätigkeit eine gesetzlich geregelte Ausbildung, staatliche Zulassung oder spezifische Berufsethik voraus?
- Erfordert die Entscheidung normative Abwägung, Kontextsensibilität oder die Übernahme persönlicher Verantwortung?
- Spielen Fähigkeiten wie Empathie, kommunikative Kompetenz oder situatives Urteil eine zentrale Rolle?

Darüber hinaus stellt sich die Frage, ob bestimmte Fähigkeiten strukturell an menschliche Entscheidungsträger gebunden sind. Dies kann etwa der Fall sein, wenn Verantwortungsübernahme rechtlich oder institutionell personengebunden ist, bestimmte Aufgaben aufgrund ihrer Natur nicht delegierbar sind oder wenn normative Begründung und Rechtfertigung integraler Bestandteil der Entscheidung sind.
Die Fragen dieses Abschnitts folgen keiner ordinalen Logik. Sie erfassen unterschiedliche Indikatoren dafür, dass eine Tätigkeit strukturell auf menschliche Urteilskompetenz angewiesen ist. Entscheidend ist daher nicht die Anzahl bejahter Punkte, sondern die qualitative Bedeutung der jeweiligen Merkmale für die konkrete Entscheidungssituation.
Wenn zentrale Elemente der Tätigkeit in normativer Abwägung, situativem Urteil, personengebundener Verantwortung oder rechtfertigender Begründung bestehen, spricht dies regelmäßig dafür, dass die Aufgabe nur begrenzt algorithmisch vorstrukturierbar ist. In solchen Konstellationen markieren bereits die Anforderungen der Tätigkeit selbst eine Grenze sinnvoller technischer Unterstützung.

dd. Struktur der Entscheidungsfindung

Dieser Abschnitt operationalisiert die Komplexität des Entscheidungsfeldes. Er prüft, ob eine Entscheidung in einem stabilen, wiederkehrenden und regelgeleiteten Umfeld stattfindet oder ob sie von Unsicherheit, situativen Besonderheiten und schwer explizierbaren Erwägungen geprägt ist. Je stärker Entscheidungen auf konsistent anwendbaren Regeln beruhen und je stabiler die Entscheidungsumgebung ist, desto eher liegt ein formalisiertes Entscheidungsfeld vor, in dem technische Unterstützung grundsätzlich anschlussfähig sein kann. Seltene, datenarme oder stark kontextabhängige Konstellationen sprechen dagegen für ein hohes Maß an Komplexität.
Der Abschnitt gibt zugleich Hinweise auf den angemessenen Automatisierungsgrad. Technische Systeme können Entscheidungsprozesse nur dort sinnvoll strukturieren oder unterstützen, wo die zugrunde liegenden Entscheidungslogiken hinreichend explizierbar und operationalisierbar sind. Wo dies nicht der Fall ist, steigt das Risiko,

dass technische Unterstützung relevante Kontextfaktoren ausblendet oder implizite Modellannahmen an die Stelle fachlicher Urteilskraft treten.
Zu prüfen ist etwa:

- Ob Entscheidungen überwiegend auf Erfahrungswissen, Intuition, impliziten Mustern oder expliziten Entscheidungsregeln beruhen.
- Ob die tatsächlich verwendeten Entscheidungsregeln dokumentiert, explizierbar und konsistent anwendbar sind.
- Ob bestimmte Informationsquellen systematisch ignoriert, übergewichtet oder je nach Situation unterschiedlich berücksichtigt werden.

Ebenso relevant ist die Struktur der Entscheidungsumgebung:

- Ist die Entscheidungssituation relativ stabil oder von hoher Unsicherheit geprägt?
- Handelt es sich um wiederkehrende Fälle oder um seltene Einzelfälle?
- Ist das Entscheidungsfeld datenreich oder datenarm?

Zu prüfen ist außerdem, welche Bewertungsstruktur der betreffenden Entscheidung zugrunde liegt:

- Existiert ein klar definierter Bewertungsmaßstab?
- Existiert ein objektivierbares Zielkriterium?
- Wie groß ist der Ermessensspielraum?
- Sind Grundrechte oder andere besonders schützenswerte Güter betroffen?

Die Fragen dieses Abschnitts folgen teilweise einer graduellen Logik. Je stabiler die Entscheidungsumgebung, je häufiger vergleichbare Fälle auftreten und je klarer Entscheidungsregeln dokumentiert und konsistent angewendet werden können, desto eher liegt ein formalisiertes Entscheidungsfeld vor. Umgekehrt sprechen seltene, datenarme und stark kontextabhängige Entscheidungssituationen für ein hohes Maß an Komplexität. Andere Fragen prüfen dagegen strukturelle Voraussetzungen der Entscheidungsmodellierung, etwa das Vorliegen klarer Bewertungsmaßstäbe oder objektivierbarer Zielkriterien.

ee. Entscheidungsvarianz und Bias-Strukturen

Die Fragen zu Entscheidungsvarianz und möglichen Bias-Strukturen erfüllen in der Kontextanalyse eine doppelte Funktion. Zum einen geben sie Hinweise darauf, ob überhaupt ein Problem besteht, das durch den Einsatz eines KI-Systems adressiert werden kann. Zum anderen liefern sie zusätzliche Anhaltspunkte für die Komplexität des Entscheidungsfeldes. Hohe inter- oder intrapersonelle Varianz kann auf Unsicherheit, fehlende Standardisierung oder schwer explizierbare Abwägungsprozesse hinweisen, sie kann jedoch ebenso Ausdruck verzerrter oder qualitativ unzureichender Entscheidungsroutinen sein.
Für die Kontextanalyse ist daher entscheidend, diese Konstellationen zu unterscheiden. Varianz ist nicht per se ein Problem. In manchen Bereichen spiegelt sie legitime Einzelfallsensibilität wider; in anderen weist sie auf korrigierbare Inkonsistenzen oder Bias-Strukturen hin. Der Abschnitt ist deshalb primär diagnostisch angelegt, er soll klären, ob ein KI-System tatsächlich ein bestehendes Problem adressiert oder ob eine Reduktion von Varianz mit dem Verlust sachlich gebotener Differenzierung verbunden wäre.

Hinsichtlich intrapersoneller Inkonsistenz ist zu prüfen:

- Führen identische oder sehr ähnliche Fälle bei derselben Person zu unterschiedlichen Entscheidungen?
- Variieren Entscheidungen abhängig von situativen Faktoren wie Zeitdruck oder Stress, von irrelevanten Kontextinformationen oder von Reihenfolgeeffekten?
- Ist Entscheidungsvarianz messbar dokumentiert?

Interpersonelle Varianz:

- Treffen unterschiedliche Akteure in vergleichbaren Fällen systematisch unterschiedliche Entscheidungen?
- Existieren institutionelle Leitlinien oder standardisierte Bewertungsmaßstäbe?
- Ist beobachtete Varianz Ausdruck gewollten Ermessensspielraums oder unbeabsichtigter Inkonsistenz?

Bias-Strukturen:

- Korrelieren Entscheidungen systematisch mit irrelevanten personenbezogenen Merkmalen?
- Korrelieren Entscheidungen mit sozialen, ethnischen oder geschlechtsspezifischen Variablen?
- Werden Proxy-Variablen faktisch verwendet?
- Bestehen historische Verzerrungen im Entscheidungsfeld?
- Sind Korrekturmechanismen vorhanden?

Bei der Beantwortung dieser Fragen ist methodische Vorsicht geboten. Belastbare Daten über Entscheidungsvarianz oder Bias-Strukturen sind in vielen institutionellen Kontexten nur begrenzt verfügbar: Interne Entscheidungsdaten sind häufig schwer zugänglich, unvollständig dokumentiert oder aufgrund geringer Fallzahlen nur eingeschränkt repräsentativ. Die Beantwortung der Fragen kann daher nicht immer ausschließlich auf unmittelbarer empirischer Evidenz beruhen. In solchen Fällen können ergänzend Ergebnisse aus empirischer Forschung in vergleichbaren Entscheidungsdomänen sowie die Einschätzungen erfahrener Praktiker herangezogen werden. Diese dienen nicht als definitive Evidenz, sondern als heuristische Orientierung für die Einordnung möglicher Inkonsistenz- oder Bias-Risiken.

ff. Algorithmisierbarkeit der Entscheidung

Dieser Abschnitt prüft, ob eine Entscheidung strukturell algorithmischer Unterstützung zugänglich ist. Maßgeblich ist dabei nicht die abstrakte Leistungsfähigkeit technischer Systeme, sondern ob sich die entscheidungsrelevanten Merkmale, Gewichtungen und Beziehungen in einer Weise formalisieren lassen, die dem tatsächlichen Charakter der Aufgabe entspricht. Regelbasierte, datengetriebene oder prognostische Teilaufgaben sind in der Regel eher algorithmisierbar als Abwägungsentscheidungen mit offenen normativen Bewertungsmaßstäben.
Für sprachlich-generative Teilaufgaben gilt eine andere Logik der Algorithmisierbarkeit: Maßgeblich ist hier nicht die Stabilität von Prognose- oder Kausalbeziehungen, sondern die Zuverlässigkeit, mit der ein System vorgegebene Quellinformationen treu wiedergibt oder zusammenfasst, ohne inhaltlich über sie

hinauszugehen (Konfabulation). Auch generative Teilaufgaben können damit algorithmisierbar sein, ohne klassifikatorisch oder prognostisch zu sein.
Für die Kontextanalyse ist dies zentral, weil hier die Verbindung zwischen Komplexität und Automatisierungsgrad sichtbar wird. Eine Entscheidung, die strukturell nur begrenzt formal beschreibbar ist, setzt einer sinnvollen technischen Unterstützung entsprechende Grenzen, und zwar unabhängig davon, was ein System im Einzelfall leisten kann. Eine weitergehende technische Vorstrukturierung birgt dann das Risiko, relevante Elemente menschlicher Urteilsbildung zu verdrängen, ohne sie angemessen ersetzen zu können: Algorithmische Empfehlungen werden faktisch zum Referenzpunkt, ohne dass die ihnen zugrunde liegenden Formalisierungen dem normativen Gehalt der Aufgabe gerecht werden.
Zu klären ist insbesondere:

- Handelt es sich um eine überwiegend regelbasierte, datengetriebene, prognostische oder klassifikatorische Entscheidung?
- Lassen sich relevante Merkmale, Gewichtungen und Entscheidungskriterien formal beschreiben und operationalisieren?
- Existieren stabile Prognose- oder Kausalbeziehungen, auf denen ein technisches System sinnvoll aufbauen kann?
- Handelt es sich um eine überwiegend sprachlich-generative Aufgabe (Zusammenfassung, Textvergleich, Formulierungsunterstützung)? Lässt sich die Zuverlässigkeit der Quellwiedergabe für den relevanten Aufgabentyp belegen?

Die Antworten erlauben eine heuristische Einordnung entlang einer graduellen Logik. Je klarer sich Entscheidungsregeln, relevante Merkmale und deren Gewichtungen formalisieren lassen, je stabiler die zugrunde liegenden Prognose- oder Kausalbeziehungen sind und je stärker die Aufgabe regelbasierten, klassifikatorischen oder prognostischen Strukturen folgt, desto eher ist eine technische Unterstützung grundsätzlich anschlussfähig. Umgekehrt spricht eine Situation, in der zentrale Entscheidungsmerkmale nicht zuverlässig operationalisierbar sind, kausale oder prognostische Beziehungen instabil bleiben oder normative Abwägungen den Kern der Entscheidung bilden, für eine geringe Algorithmisierbarkeit. In solchen Fällen können technische Systeme allenfalls unterstützende Teilfunktionen übernehmen, ohne den Entscheidungsprozess selbst strukturell vorzuformen.
Für generative Teilaufgaben gilt Entsprechendes: Je zuverlässiger belegt ist, dass ein System Quellinformationen treu wiedergibt, ohne zu konfabulieren, desto eher ist auch eine sprachlich-generative Unterstützung anschlussfähig; fehlt ein solcher Beleg, gilt dieselbe Zurückhaltung wie bei instabilen Prognose- oder Kausalbeziehungen.

gg. Ziel des KI-Einsatzes

Die vorangegangenen Teilabschnitte haben die Entscheidungssituation entlang ihrer normativen, strukturellen und empirischen Dimensionen analysiert. Die Ausgangsfrage war, welches konkrete Problem durch den Einsatz eines KI-Systems gelöst werden soll und welcher Teil des Entscheidungsprozesses davon tatsächlich betroffen ist.
Die Platzierung am Ende ist methodisch begründet. In der Praxis werden bei der Planung technischer Systeme einzelne Probleme häufig überproportional gewichtet, während andere Aspekte der Entscheidungssituation aus dem Blick geraten. In der

Entscheidungsforschung wird diese Verzerrung als Focusing Illusion beschrieben.[32] Wer die Zielfrage zu Beginn stellt, läuft Gefahr, die Kontextanalyse auf das vorab formulierte Ziel hin zu verengen. Wer sie am Ende stellt, kann prüfen, ob das ursprünglich identifizierte Problem nach der Gesamtanalyse noch dieselbe Gestalt hat oder ob die Analyse neue Problemdimensionen sichtbar gemacht hat, die eine Neujustierung erfordern.
Die Reflexion des Ziels schützt damit vor zwei gegenläufigen Fehlermustern. Zum einen verhindert sie, dass KI zum Selbstzweck wird. Die Entscheidung für oder gegen ein System sollte nicht primär von technischer Verfügbarkeit abhängen, sondern von der Passung zwischen Problemstruktur, technischer Interventionsfähigkeit und den erwartbaren Risiken. Ferner verhindert sie eine vorschnelle Ablehnung. Wenn die vorangegangene Analyse ein klar identifizierbares Problem und eine geeignete technische Interventionsmöglichkeit ergibt, wäre der Verzicht auf KI-Unterstützung seinerseits begründungspflichtig. Dies schließt an die Symmetriebedingung an: Opportunitätskosten der Nichtnutzung sind nur dann analytisch tragfähig, wenn das adressierte Problem empirisch belegt und das Ziel des Eingriffs präzise definiert ist.
Zu prüfen ist daher abschließend, welches Problem durch den KI-Einsatz konkret adressiert werden soll. Ist dieses Problem durch die vorangegangene Analyse empirisch gestützt oder handelt es sich um eine Annahme, die noch zu belegen wäre? Ist das angestrebte Ziel mit dem normativen Charakter der Entscheidung, dem Komplexitätsgrad des Feldes und der Algorithmisierbarkeit der Aufgabe vereinbar? Und schließlich: Überwiegen die erwartbaren Gewinne an Effizienz oder Entscheidungsqualität die identifizierten Risiken – einschließlich jener systemischen Risiken, die erst bei breiter Implementation sichtbar werden? Auf welcher Akteursebene entstehen die erwarteten Gewinne, und auf welcher Ebene werden die Risiken verteilt? Sind Effizienzgewinne auf der sekundären Ebene – etwa Entlastung von Gerichten – mit Qualitäts- oder Fairnessverlusten auf der primären verbunden? Wurde bei der Zieldefinition die Risikotoleranz der primär Betroffenen einbezogen oder implizit durch institutionell distanzierte Akteure ersetzt?

d. Menschliche Kontrolle als institutionelle Leistung

Die Anwenderfragen der Abschnitte C.I.4.c.aa bis C.I.4.c.gg haben die Entscheidungssituation entlang der drei Risikodimensionen – Normativität, Automatisierungsgrad und Komplexität – systematisch eingeordnet. Sie beschreiben, welche Güter berührt werden, wie eingriffsintensiv und reversibel Fehlentscheidungen sind, in welchem Maß der Entscheidungsprozess formalisierbar ist und welche Teilaufgaben sich algorithmisch unterstützen lassen. Damit ergeben ihre Antworten das Risikoprofil einer Anwendung. Für die Verdichtung zu einer Risikostufe können die zentralen Risikodimensionen aus C.I.4.b dienen: Ein empfehlender oder vorstrukturierender Einsatz ist mindestens als hohes Risiko einzustufen, wenn die Tätigkeit hoch normativ ist und entweder das Entscheidungsfeld komplex ist oder die Entscheidungsfolgen faktisch irreversibel sind. Ein informierender Einsatz in einem solchen Kontext ist mindestens als moderates Risiko einzustufen, weil auch die Auswahl und Aufbereitung von Informationen die Entscheidung vorprägen kann. Ein niedriges Risikoprofil setzt voraus, dass die Tätigkeit überwiegend administrativ, das System informierend und die Entscheidungsfolgen reversibel sind. Alle übrigen Konstellationen sind als moderat einzustufen, sofern die Gesamtbewertung nicht begründet zu einem anderen Ergebnis führt.

[32] *Schkade/Kahneman*, Does Living in California Make People Happy? A Focusing Illusion in Judgments of Life Satisfaction, in: Psychological Science 9 (1998), 340.

Dieses Profil ist damit Ausgangspunkt der Analyse. Es bestimmt, welche Risikoszenarien plausibel sind und welche institutionellen Bedingungen ihre Realisierung begünstigen oder hemmen. Zentral ist dabei die Frage menschlicher Kontrolle.
Menschliche Kontrolle ist dabei kein ausschließlich aus der KI-Verordnung abgeleitetes Konzept. Die CEPEJ (European Commission for the Efficiency of Justice) des Europarats hat bereits 2018 mit der European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems fünf Prinzipien formuliert – Achtung der Grundrechte, Diskriminierungsfreiheit, Qualität und Sicherheit, Transparenz, Unparteilichkeit und Fairness sowie Nutzerkontrolle –, die menschliche Kontrolle als eigenständiges, bereichsspezifisches Prinzip für den Justizkontext verankern.[33] In ihren im Dezember 2025 verabschiedeten Guidelines on the Use of Generative AI for Courts konkretisiert die CEPEJ dies für generative Systeme: Die richterliche Gewalt verbleibt exklusiv bei den Gerichten, der Zugang zu einem menschlichen Richter ist stets gewährleistet, Outputs generativer Systeme sind niemals bindend, und ihr Einsatz muss transparent erfolgen.[34] Diese Prinzipien decken sich in ihrer Struktur mit den im Folgenden entwickelten Aufsichtsbedingungen und bestätigen aus bereichsspezifischer Perspektive, was die KI-Verordnung horizontal für alle Hochrisikosysteme fordert.

aa. Grenzen des Human-Oversight-Modells

Regulatorische Ansätze gehen häufig davon aus, dass die Einbindung eines menschlichen Entscheidungsträgers eine wirksame Kontrolle algorithmischer Systeme gewährleistet. Empirische und theoretische Arbeiten stellen diese Annahme in Frage. Green[35] zeigt, dass entsprechende Modelle selten prüfen, ob menschliche Entscheider unter realen Arbeitsbedingungen tatsächlich in der Lage sind, algorithmische Ergebnisse kritisch zu überprüfen. Für den Justizkontext spezifisch bestätigt dies Banks[36] mit Blick auf die Aufsichtspflichten der KI-Verordnung. Menschliche Beteiligung kann unter solchen Bedingungen symbolischen Charakter annehmen und algorithmische Entscheidungen nachträglich legitimieren, statt sie zu kontrollieren.
Daraus folgt, dass menschliche Beteiligung nicht als gegebenes Sicherheitsmerkmal behandelt werden darf. Entscheidend ist, unter welchen Bedingungen Kontrolle tatsächlich wirksam wird.

[33] *European Commission for the Efficiency of Justice (CEPEJ)*, European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment, 2018, abrufbar unter https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c (Abruf am 24.04.2026, 14:58 Uhr).
[34] *European Commission for the Efficiency of Justice (CEPEJ)*, Guidelines on the use of generative AI for courts, CEPEJ(2025)18Final, 2025, abrufbar unter https://rm.coe.int/cepej-2025-18final-en-draft-guidelines-on-the-use-of-generative-ai-for/48802a4ad1 (Abruf am 25.10.2026, 14:29 Uhr).
[35] *Green*, The flaws of policies requiring human oversight of government algorithms, Computer Law & Security Review 45 (2022), Art. 105681.
[36] *Banks*, Judges-in-the-loop? Judicial involvement in human oversight of high-risk decision support systems under the EU AI Act, International Journal of Law and Information Technology 34 (2026), eaag001.

bb. Vier Bedingungen wirksamer Aufsicht

Sterz et al.[37] schlagen vor, die Wirksamkeit menschlicher Kontrolle entlang von vier Bedingungen zu analysieren:

- **Kausale Eingriffsmacht**: Die aufsichtsführende Person muss technisch und organisatorisch in der Lage sein, Systemoutputs zu verändern oder unabhängig vom System zu entscheiden.
- **Epistemischer Zugang**: Sie benötigt hinreichende Transparenz über Datenbasis, Modellannahmen und Unsicherheiten, um Ergebnisse kritisch einordnen zu können.
- **Selbststeuerungsfähigkeit**: Zeitdruck, Arbeitsbelastung oder kognitive Überforderung können vorhandene Kontrollmöglichkeiten faktisch entwerten.
- **Passende Absichten:** Die aufsichtsführende Person muss die Absicht haben, die durch den KI-Einsatz entstehenden Risiken zu mindern.

Diese vier Bedingungen verweisen auf drei wechselwirkende Einflussbereiche der Mensch-System-Interaktion: erstens technische Designentscheidungen (Interface-Gestaltung, Darstellung von Unsicherheit, Interventionsmöglichkeiten), zweitens individuelle Faktoren der Entscheidungsträger (Erfahrung, Fachwissen, Motivation) und drittens organisationale Rahmenbedingungen (Arbeitsbelastung, Zeitdruck, Erwartungshaltungen). Risiken entstehen häufig nicht aus einzelnen Faktoren, sondern aus deren Zusammenspiel. Der Bericht liest diese Bedingungen institutionell: Auch Absichten sind nicht nur individuelle Haltungen, sondern können durch organisationale Anreizstrukturen verschoben werden, etwa wenn Effizienzerwartungen die Absicht, Risiken zu mindern, hinter das Ziel zügiger Erledigung zurücktreten lassen (vgl. Szenario 3).

cc. Meaningful Human Intervention

Eng verbunden mit den Aufsichtsbedingungen ist das Konzept der Meaningful Human Intervention, das aus dem Datenschutzrecht stammt: Nach den Leitlinien zu Art. 22 DSGVO muss menschliche Beteiligung an einer Entscheidung bedeutsam und nicht bloß eine symbolische Geste sein.[38] Menschliche Beteiligung ist nur dann substantiell, wenn Entscheidungsträger und Entscheidungsträgerinnen zusätzliche Informationen einbeziehen, Besonderheiten eines Falles berücksichtigen und algorithmische Ergebnisse korrigieren oder ergänzen können. Fehlt dieser Handlungsspielraum, etwa durch technische Gestaltung, zeitliche Belastung oder institutionelle Erwartungshaltungen, besteht die Gefahr, dass menschliche Beteiligung zur formalen Bestätigung algorithmischer Ergebnisse wird. Dies ist das Muster, das als faktische Begründungslast bei Abweichung von Systemempfehlungen bereits beschrieben wurde, hier unter dem Gesichtspunkt seiner institutionellen Entstehungsbedingungen.

[37] *Sterz/Baum/Biewer/Hermanns/Lauber-Rönsberg/Meinel/Langer*, On the Quest for Effectiveness in Human Oversight: Interdisciplinary Perspectives, 2024, abrufbar unter https://doi.org/10.48550/arXiv.2404.04059 (Abruf am 22.10.2026, 20:38 Uhr), S. 4f.

[38] *Artikel-29-Datenschutzgruppe*, WP 251 rev.01, Guidelines on automated individual decision-making and profiling for the purposes of Regulation 2016/679, abrufbar unter https://www.edpb.europa.eu/documents/guideline/automated-decision-making-and-profiling_en (Abruf am 25.10.2026, 14:23 Uhr).

dd. Typische Risikopfade

Besonders risikoreiche Konstellationen entstehen dort, wo hohe normative Bedeutung, ein komplexes Entscheidungsfeld und ein hoher Grad technischer Vorstrukturierung zusammentreffen. Aus den Aufsichtsbedingungen und den drei Einflussbereichen lassen sich vier typische Risikopfade rekonstruieren:

- **Epistemische Überbewertung**: Algorithmische Ergebnisse werden als objektiver oder zuverlässiger wahrgenommen, als es ihre epistemische Grundlage rechtfertigt.
- **Eingeschränkte Intervention**: Override-Möglichkeiten existieren formal, werden aber durch asymmetrischen Begründungsaufwand oder organisatorischen Druck faktisch entwertet.
- **Institutionelle Standardisierung**: Systemempfehlungen werden schrittweise zum Referenzpunkt, sodass Abweichungen erklärungsbedürftig werden, ohne dass dies formell vorgesehen ist.
- **Fehlinterpretation**: Algorithmische Ausgaben – probabilistische Prognosen ebenso wie generierte Texte – werden als geprüfte, deterministische Aussagen über den Einzelfall interpretiert, obwohl ihre epistemische Grundlage dies nicht trägt.

Diese Pfade entstehen typischerweise aus dem Zusammenspiel technischer Systemgestaltung, individueller Entscheidungsheuristiken und organisationaler Rahmenbedingungen. Der folgende Abschnitt illustriert sie exemplarisch.

e. Qualitative Risikoanalyse: Szenarien

Die folgenden Szenarien illustrieren typische Risikokonfigurationen, die sich aus den in dd. identifizierten Risikopfaden ergeben. Sie stellen keine Prognosen über konkrete Systeme dar, sondern rekonstruieren plausible Wirkzusammenhänge zwischen Systemgestaltung, Entscheidungspraxis und institutionellen Rahmenbedingungen. Ausgangsbedingung, Interaktionsdynamik, Verhaltensanpassung, Ergebnis und spezifisches Risiko geben dabei die Struktur vor. Die Szenarien unterscheiden sich hinsichtlich des primären Risikopfades und der jeweils verletzten Aufsichtsbedingung.

aa. Szenario 1: Automation Bias durch epistemische Überbewertung

Primärer Risikopfad: Epistemische Überbewertung. Verletzte Aufsichtsbedingung: Epistemischer Zugang.
Ein algorithmisches System generiert auf Grundlage statistischer Modelle Empfehlungen für richterliche Entscheidungen, etwa Risikoscores oder Strafempfehlungen. Die Empfehlung wird als numerischer Score dargestellt; Informationen über Modellunsicherheit, Datenbasis oder Prognosegrenzen sind nur eingeschränkt sichtbar. Diese Gestaltung reproduziert die Gefahr der Scheingenauigkeit auf der Ebene der Nutzerinteraktion.
Bei generativ arbeitenden Systemen tritt dasselbe Muster in sprachlicher statt numerischer Form auf: Ein flüssig formulierter, selbstsicher wirkender Textentwurf erzeugt einen vergleichbaren Anschein von Autorität wie ein numerischer Score, ohne dass die zugrundeliegende Unsicherheit sprachlich markiert wird.
Die visuelle Darstellung erzeugt den Eindruck einer objektiven, datenbasierten Bewertung. Die algorithmische Empfehlung tritt als epistemisch privilegierte

Informationsquelle auf und untergräbt die epistemische Zugangsbedingung wirksamer Aufsicht. Der menschliche Entscheider oder die menschliche Entscheiderin gewichtet den Systemoutput stärker als eigene Beobachtungen oder fallbezogene Besonderheiten. Abweichungen erscheinen kognitiv erklärungsbedürftig.
Ergebnis: Algorithmische Empfehlungen werden systematisch häufiger übernommen, auch wenn eine eigenständige normative Abwägung zu einem anderen Ergebnis führen könnte.
Risiko: Normative Bewertungsentscheidungen verschieben sich faktisch auf das algorithmische System. Die richterliche Abwägung wird durch die epistemische Autorität des Systems vorstrukturiert. Meaningful Human Intervention wird zur formalen Bestätigung.[39]

bb. Szenario 2: Institutionelles Handicapping effektiver Kontrolle

Primärer Risikopfad: Eingeschränkte Intervention. Verletzte Aufsichtsbedingung: Kausale Eingriffsmacht und passende Absichten.
Während Szenario 1 die kognitive Ebene des einzelnen Entscheiders betrifft, adressiert dieses Szenario die institutionelle Ebene. Das System sieht formal vor, dass Entscheidungsträger Empfehlungen überprüfen und übersteuern können. Eine Abweichung ist technisch möglich, erfordert jedoch zusätzliche Dokumentation oder gesonderte Begründung. Die kausale Eingriffsmacht ist formal gegeben, ihre Ausübung jedoch mit asymmetrischem Aufwand verbunden.
Institutionelle Rahmenbedingungen, wie etwa hoher Arbeitsdruck, knappe Entscheidungszeiten, Effizienzvorgaben, verstärken diesen Effekt. Die versteckten Kosten einer begründeten Abweichung werden organisatorisch nicht aufgefangen. Anwender folgen Systemempfehlungen häufiger, um zusätzlichen Begründungsaufwand zu vermeiden. Damit entsteht die Gefahr, dass nicht die Übernahme erklärungsbedürftig wird, sondern die Abweichung.
Ergebnis: Override-Möglichkeiten existieren formal, werden jedoch selten genutzt. Das System ist de facto bindend, ohne de jure bindend zu sein.
Risiko: Die Entscheidung wird faktisch automatisiert, obwohl institutionell von einer menschlichen Letztentscheidung ausgegangen wird. Menschliche Beteiligung erfüllt primär eine legitimatorische Funktion. Dieses Muster ist besonders folgenreich, weil es nach außen unsichtbar bleibt: Formal sind alle Kontrollanforderungen erfüllt, faktisch ist Meaningful Human Intervention auf eine Zustimmungsgeste reduziert.[40]

cc. Szenario 3: Schleichende Standardisierung von Entscheidungspraktiken

Primärer Risikopfad: Institutionelle Standardisierung. Verletzte Aufsichtsbedingung: Unterminieren passender Absichten durch organisationale Anreizstrukturen.
Dieses Szenario betrifft die systemische Ebene und realisiert den Übergang vom idiosynkratischen zum systemischen Risiko. Ein algorithmisches System wird als freiwillige Entscheidungshilfe eingeführt. Es generiert Empfehlungen, die auf statistischen Mustern historischer Entscheidungen beruhen, und modelliert damit

[39] *Sterz/Baum/Biewer/Hermanns/Lauber-Rönsberg/Meinel/Langer*, On the Quest for Effectiveness in Human Oversight: Interdisciplinary Perspectives, 2024, abrufbar unter https://doi.org/10.48550/arXiv.2404.04059 (Abruf am 22.10.2026, 20:38 Uhr), S. 4ff.; *Green*, The flaws of policies requiring human oversight of government algorithms, Computer Law & Security Review 45 (2022), Art. 105681.
[40] *Green*, The flaws of policies requiring human oversight of government algorithms, Computer Law & Security Review 45 (2022), Art. 105681.

implizit, was als normale Entscheidung gilt, ohne diesen Normalitätsbegriff als solchen zu kennzeichnen.
Mit zunehmender Nutzung entstehen damit institutionelle Erwartungen an konsistente Entscheidungspraktiken. Abweichungen von Systemempfehlungen werden zunehmend erklärungsbedürftig. Entscheidungsträger orientieren sich stärker an den algorithmischen Empfehlungen, um institutionelle Erwartungen zu erfüllen oder Kritik an inkonsistenten Entscheidungen zu vermeiden. Dieser Prozess ist graduell und weitgehend unsichtbar.
Ergebnis: Die Varianz richterlicher Entscheidungen nimmt ab, auch in Bereichen, in denen unterschiedliche Bewertungen normativ zulässig oder geboten wären.
Risiko: Ein lokal begrenztes Werkzeug erzeugt durch Skalierung und institutionelle Kopplung strukturelle Verzerrungen, die sich über viele Verfahren reproduzieren. Kurzfristige Varianzreduktion kann mit dem Verlust sachlich gebotener Differenzierung erkauft werden. Ob das Ergebnis normativ akzeptabel ist, hängt davon ab, ob die systemimplizierte Normalverteilung dem entspricht, was das Recht in diesen Fällen tatsächlich fordert.

dd. Szenario 4: Fehlinterpretation der epistemischen Grundlage algorithmischer Ausgaben

Primärer Risikopfad: Fehlinterpretation. Verletzte Aufsichtsbedingung: Epistemischer Zugang und Selbststeuerungsfähigkeit.
Dieses Szenario betrifft die epistemische Ebene im engeren Sinne: Der Output eines Systems wird für etwas gehalten, das er nicht ist. Wie sich das zeigt, hängt vom Systemtyp ab, das zugrundeliegende Muster ist jedoch dasselbe.
Bei klassifikatorisch-prognostischen Systemen liefert das System probabilistische Einschätzungen über zukünftige Ereignisse, etwa Rückfallwahrscheinlichkeiten. Diese sind ihrem Wesen nach Aussagen über Gruppen vergleichbarer Fälle, nicht über den Einzelfall. Modellannahmen, Unsicherheitsintervalle und Datenabhängigkeiten werden nur begrenzt kommuniziert. Die Präsentation eines numerischen Scores erzeugt den Eindruck von Messbarkeit, wo tatsächlich Unsicherheit besteht.
Bei generativ arbeitenden Systemen tritt strukturell dasselbe Muster in anderer Form auf: Ein sprachlich flüssiger, geschlossen wirkender Text erzeugt den Eindruck einer geprüften, belegten Aussage, obwohl er tatsächlich eine statistische Fortsetzung wahrscheinlicher Wortfolgen ist, die von den zugrundeliegenden Quellen abweichen kann (Konfabulation). Anders als bei einem numerischen Score liegt hier keine explizite Unsicherheitsangabe vor, die unterlaufen werden könnte – die sprachliche Form selbst ist die Fehlerquelle. Die Größenordnung ist dabei praktisch relevant: In einer Untersuchung führender KI-gestützter Rechercheprodukte aus den USA, die mit Retrieval-Augmented Generation arbeiten (Lexis+ AI, Westlaw AI-Assisted Research), enthielten zwischen 17 und 33 Prozent der Antworten unzutreffende oder nicht durch die angegebenen Quellen gedeckte Aussagen.[41] Die untersuchten Produktversionen stammen aus dem Jahr 2024; die Werte sind daher nicht als aktuelle Leistungsangabe zu lesen, wohl aber als Beleg dafür, dass Konfabulation auch bei juristisch spezialisierten und quellengestützten Systemen in erheblichem Umfang auftritt. Dass generative Fehlleistungen kein theoretisches Risiko sind, zeigen zahlreiche seit 2023 und verstärkt seit 2025 gerichtlich sanktionierte Fälle erfundener Zitate durch

[41] *Magesh/Surani/Dahl/Suzgun/Manning/Ho*, Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools, Journal of Empirical Legal Studies 22 (2025), 216.

generative KI in Schriftsätzen, mit Sanktionen von wenigen Tausend bis über 100.000 US-Dollar.[42]
In beiden Fällen interpretieren Anwender den Output als deterministische oder jedenfalls belastbar geprüfte Aussage oder überschätzen seine Aussagekraft für die konkrete Entscheidungssituation. Diese Fehlinterpretation ist nicht primär ein Kompetenzdefizit, sondern ein durch die Systemgestaltung induziertes Wahrnehmungsmuster.
Ergebnis: Algorithmische Ausgaben werden überinterpretiert oder auf Entscheidungssituationen angewendet, für die sie nur begrenzte Aussagekraft besitzen.
Risiko: Entscheidungen beruhen auf einer Fehlinterpretation der epistemischen Reichweite algorithmischer Modelle. Substantielle menschliche Kontrolle würde voraussetzen, dass Entscheidungsträger die epistemischen Grenzen des Systems einschätzen und ihre Entscheidung entsprechend kalibrieren können. Wo die Systemgestaltung diese Einschätzung strukturell erschwert, ist die Kontrollbedingung auch dann nicht erfüllt, wenn formal eine menschliche Letztentscheidung vorliegt.

f. Reflexions- und Evaluationsfragen für Systemdesigner

Die Anwenderperspektive hat gezeigt, dass Risiken beim KI-Einsatz in der richterlichen Entscheidungsfindung maßgeblich durch die Struktur der Entscheidungssituation bestimmt werden. Die Perspektive der Systemdesigner setzt an einem anderen Punkt an: Sie betrifft Akteure, die durch technische und gestalterische Entscheidungen die Bedingungen schaffen, unter denen Anwender mit einem System interagieren. Ihre Entscheidungen prägen, welche Informationen sichtbar werden, wie stark Empfehlungen strukturiert sind und welche Eingriffsmöglichkeiten bestehen.
Für eine verantwortbare Implementierung ist nicht nur die Leistungsfähigkeit eines Modells relevant, sondern auch die Gestaltung der Schnittstelle zwischen System und menschlicher Entscheidung. Diese Schnittstelle beeinflusst Wahrnehmung, Aufmerksamkeit und Entscheidungslogik und kann die in C.I.4.d.dd identifizierten Risikopfade begünstigen oder hemmen.
Die folgenden Fragen sind entlang der drei Risikodimensionen aus C.I.4.b strukturiert: Die Abschnitte zur Interface-Gestaltung und Verhaltenslenkung betreffen primär den Automatisierungsgrad, da sie bestimmen, wie stark das System den Entscheidungsprozess faktisch vorstrukturiert. Die Abschnitte zu Funktionsweise und Eingriffsmöglichkeiten operationalisieren die Schnittstelle zwischen Automatisierung und Normativität. Die Abschnitte zur epistemischen Transparenz und Risikokommunikation betreffen vor allem die Aufsichtsbedingung des epistemischen Zugangs: Sie bestimmen, ob Anwender die Grenzen des Systems gegenüber der Komplexität ihres Entscheidungsfeldes einschätzen können.

aa. Interface-Gestaltung (Steuerungsarchitektur)

Die Gestaltung des Interfaces bestimmt maßgeblich, wie Nutzer Informationen wahrnehmen und verarbeiten. Sie ist der primäre technische Hebel, über den Automation Bias, epistemische Überbewertung und Fehlinterpretation algorithmischer

[42] *Norton Rose Fulbright*, AI in litigation: Update on Gen AI sanctions in 2026, 2026, abrufbar unter https://www.nortonrosefulbright.com/en-us/knowledge/publications/792d8bf3/ai-in-litigation-update-on-gen-ai-sanctions-in-2026 (Abruf am 25.10.2026).

Prognosen (Szenarien 1 und 4) entstehen oder verhindert werden können. Zu prüfen ist:

- Ist das Interface klar strukturiert, mit nachvollziehbaren Hierarchien und reduzierter Komplexität? Oder präsentiert es eine Vielzahl von Scores und Kennzahlen, die kognitive Überlastung erzeugen können?
- Werden Unsicherheiten sichtbar gemacht, etwa durch Konfidenzintervalle oder durch Offenlegung der Faktoren, die die Unsicherheit begründen? Oder erscheinen Systemoutputs als scheinbar eindeutige Ergebnisse?
- Wird der Output klar als Empfehlung gekennzeichnet? Oder entsteht der Eindruck einer faktischen Entscheidung?
- Werden von einem System generierte Textinhalte visuell von menschlich verfassten oder verifizierten Inhalten unterschieden, und ist erkennbar, auf welche Quellen sich ein generierter Text stützt?

bb. Verhaltenslenkung und Choice Architecture

Interface-Design beeinflusst nicht nur Informationsverarbeitung, sondern auch Entscheidungsverhalten. In der Entscheidungsforschung wird dies als Choice Architecture bezeichnet. Im Justizkontext ist diese Dimension besonders relevant, weil Verhaltenslenkung durch Gestaltung – im Gegensatz zu expliziten Vorgaben – institutionell unsichtbar bleibt und formaler Kontrolle ohne faktische Wirksamkeit Vorschub leisten kann (Szenario 2). Zu prüfen ist:

- Werden Default-Optionen gesetzt, etwa vorselektierte Empfehlungen? Falls ja: Sind diese sachlich begründet und dokumentiert?
- Begünstigt das Design eine unkritische Übernahme von Systemempfehlungen, etwa durch schnelle Bestätigungsoptionen statt eigenständiger Prüfung?
- Werden Nutzer über bekannte Risiken wie Automation Bias informiert?
- Erzeugen bestimmte Gestaltungselemente faktisch Verhaltenslenkung, ohne diese Wirkung transparent zu machen?

cc. Funktionsweise und Delegationsgrad

Eine zentrale Designentscheidung betrifft die Rolle des Systems im Entscheidungsprozess. Sie entspricht auf der technischen Ebene der Automatisierungsdimension und bestimmt, wie weit die Entscheidung faktisch auf das System verlagert wird. Zu klären ist:

- Handelt es sich um ein informierendes System (zusätzliche Informationen), ein strukturierendes System (Sortierung, Priorisierung), ein empfehlendes System (konkrete Handlungsvorschläge) oder ein teilautomatisiertes System?
- Welche Teilaufgaben übernimmt das System konkret: Auswahl relevanter Daten, Gewichtung von Informationen, Klassifikation, Priorisierung oder Generierung eines Entscheidungsvorschlags?

Selbst wenn formal keine automatisierte Entscheidung erfolgt, kann eine starke Vorstrukturierung durch Rankings oder Scores die Entscheidungsfreiheit faktisch einschränken. Dies ist der Mechanismus, der zu schleichender Standardisierung führt

(Szenario 3), ohne dass ein einzelner Designschritt dafür verantwortlich gemacht werden kann.

dd. Eingriffsmöglichkeiten und menschliche Kontrolle

Dieser Abschnitt operationalisiert die kausale Eingriffsmacht als Aufsichtsbedingung (C.I.4.d.bb). Zu prüfen ist:

- Können Anwender Systemoutputs korrigieren oder anpassen? Lassen sich Parameter oder Entscheidungsschritte verändern? Ist ein Override technisch möglich?
- Ist ein Override mit zusätzlichen Dokumentationspflichten verbunden? Wird er organisatorisch erschwert oder implizit sanktioniert? Werden Override-Entscheidungen systematisch ausgewertet?
- Gibt es Interventionsmöglichkeiten vor der Entscheidung, während der Verarbeitung oder erst nachträglich?

Je früher Eingriffe möglich sind, desto größer ist die tatsächliche Kontrolle über den Entscheidungsprozess. Je asymmetrischer der Aufwand zwischen Übernahme und Abweichung, desto stärker das Risiko einer institutionellen Erosion effektiver Kontrolle (Szenario 2).
Eine ergänzende Frage betrifft Eingriffsmöglichkeiten nicht der Anwender, sondern der von der Entscheidung Betroffenen selbst. Prüfarchitekturen wie das A5-Basismodul des BSI sehen für KI-Systeme allgemein eigenständige Anfechtungs- und Korrekturmechanismen vor, über die Nutzende oder betroffene Personen algorithmisch mitgeprägte Ergebnisse anfechten können.[43] Im gerichtlichen Verfahren wird diese Funktion strukturell durch das bestehende Rechtsmittelsystem übernommen: Instanzenzüge, Begründungspflichten und Verfahrensgarantien (vgl. C.I.2.b) eröffnen Verfahrensbeteiligten bereits einen Weg, eine – auch algorithmisch mitgeprägte – Entscheidung anzufechten, ohne dass ein zusätzlicher, systemseitiger Anfechtungskanal erforderlich wäre. Diese Äquivalenz ist jedoch nicht voraussetzungslos: Sie funktioniert nur, wenn die Begründung einer Entscheidung den tatsächlichen Einfluss des Systems erkennen lässt (vgl. C.I.4.f.ee) und die Anfechtung nicht durch dieselben Mechanismen erschwert wird, die bereits die Anwenderebene betreffen (etwa Begründungsasymmetrien, Szenario 2).
Zu prüfen ist daher ergänzend:

- Lässt sich aus der Entscheidungsbegründung erkennen, in welchem Umfang eine Systemempfehlung die Entscheidung mitgeprägt hat?
- Steht Verfahrensbeteiligten damit faktisch ein Weg offen, den Einfluss des Systems im Rechtsmittelverfahren zu adressieren?

ee. Epistemische Transparenz

Anwender können algorithmische Ergebnisse nur dann kritisch einordnen, wenn sie verstehen, auf welcher Grundlage diese entstanden sind und wo ihre Grenzen liegen.

[43] *Bundesamt für Sicherheit in der Informationstechnik (BSI)*, A5 – Horizontales Trustworthiness-Basismodul, Community-Draft, Version 0.9, Stand 30.06.2026, abrufbar unter https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KI/A5/A5_Horizontales-Trustworthiness-Basismodul.pdf (Abruf am 25.10.2026, 14:27 Uhr).

Damit wird die epistemische Zugangsbedingung (C.I.4.d.bb) auf der Designebene adressiert. Zu prüfen ist:

- Ist dokumentiert, welche Datenquellen verwendet, welche Zielgrößen optimiert und welche Modellklassen eingesetzt werden?
- Sind diese Informationen für Anwender verständlich aufbereitet und nicht nur technisch korrekt?
- Können Systementscheidungen fallbezogen erklärt werden? Oder existieren lediglich globale Modellbeschreibungen?

ff. Risikokommunikation

Epistemische Transparenz über die Systemlogik ist notwendig, aber nicht hinreichend. Verantwortbare Systemgestaltung erfordert darüber hinaus eine transparente Kommunikation der Grenzen und Risiken des Systems, einschließlich solcher, die erst im Zusammenspiel mit menschlichen Entscheidungsprozessen und institutionellen Rahmenbedingungen entstehen. Zu prüfen ist:

- Werden potenzielle Fehlklassifikationen, systematische Verzerrungen oder Datenabhängigkeiten offen benannt?
- Werden Risiken rein menschlicher Entscheidungsprozesse, etwa Inkonsistenz oder situative Verzerrungen, ebenfalls thematisiert? Symmetriebedingung: Einseitige Darstellung der Systemleistung ohne Benennung von Grenzen ist keine neutrale Information.
- Wird klar kommuniziert, in welchen Situationen das System sinnvoll eingesetzt werden kann und in welchen seine Nutzung ausdrücklich nicht empfohlen wird?

Verantwortbare Systemgestaltung schließt die Bereitschaft ein, die eigenen Produkte dort zu begrenzen, wo ihr Einsatz mit den in der Kontextanalyse identifizierten Risikoprofilen nicht vereinbar ist.

g. Reflexions- und Evaluationsfragen für Organisationen

Die Perspektive der Systemdesigner hat gezeigt, dass technische Gestaltungsentscheidungen die Bedingungen menschlicher Kontrolle maßgeblich prägen, sie jedoch nicht allein garantieren können. Selbst sorgfältig gestaltete Systeme können problematische Effekte entfalten, wenn die institutionellen Rahmenbedingungen ihres Einsatzes unzureichend sind. Die organisationale Perspektive setzt an diesem Punkt an, indem sie die Bedingungen adressiert, unter denen technische Systeme in der Praxis wirken.
Organisationen sind nicht nur Nutzer technischer Systeme, sondern deren institutionelle Ermöglichungsbedingung. Bezogen auf die Risikodimensionen aus C.I.4.b operationalisiert die organisationale Perspektive insbesondere die Schnittstelle zwischen Automatisierungsgrad und Normativität. Sie bestimmt, ob die in C.I.4.f adressierten technischen Kontrollmöglichkeiten im Alltag wirksam werden oder institutionell unterlaufen werden.

aa. Erwartungs- und Rollenklarheit

Eine zentrale Voraussetzung für einen verantwortbaren KI-Einsatz ist die klare Definition der Systemrolle im Entscheidungsprozess. Fehlt diese Klarheit, entstehen

implizite Erwartungen, die den Übergang von freiwilliger Nutzung zu faktischer Standardisierung begünstigen (Szenario 3). Zu prüfen ist:

- Ist eindeutig festgelegt, ob das KI-System verpflichtend, optional oder lediglich unterstützend eingesetzt wird?
- Ist klar definiert, in welchen Entscheidungssituationen es genutzt werden soll und in welchen seine Nutzung ausgeschlossen ist?
- Wer trägt die Letztverantwortung für Entscheidungen unter KI-Einsatz? Wer dokumentiert Abweichungen? Wer bewertet die Systemzuverlässigkeit im laufenden Betrieb?
- Erzeugen Leistungskennzahlen oder Effizienzvorgaben unbeabsichtigt Anreize zur bevorzugten Übernahme von Systemempfehlungen?

bb. Kommunikationsstrukturen

Wirksame menschliche Kontrolle setzt funktionierende institutionelle Kommunikationsstrukturen voraus. Ohne strukturierte Rückkanäle zwischen Anwendern, Organisationen und Systemdesignern bleiben Fehlfunktionen und unerwartete Systemeffekte unsichtbar. Zu prüfen ist:

- Existieren strukturierte Kanäle, über die Anwender Fehlfunktionen, unerwartete Ergebnisse oder mögliche Bias-Strukturen melden können?
- Werden solche Rückmeldungen systematisch ausgewertet?
- Werden Systemupdates transparent kommuniziert und deren Auswirkungen auf Entscheidungsprozesse erklärt?

Besondere Aufmerksamkeit verdienen Informationsasymmetrien zwischen technischen Entwicklern und operativen Anwendern. In vielen Organisationen verfügen Entwickler über Detailwissen zur Systemfunktionsweise, während Anwender primär mit den Outputs arbeiten. Diese Asymmetrie begünstigt die Fehlinterpretation algorithmischer Prognosen (Szenario 4). Das heißt nicht zwangsläufig, dass Anwender unkritisch sind, sondern dass ihnen womöglich die Grundlage für eine informierte Einschätzung fehlt.

cc. Umgang mit gewollter Nichtnutzung

Wenn die Entscheidung gegen den Einsatz eines Systems eine rationale Reaktion auf schwer quantifizierbare Risiken sein kann (vgl. C.I.2.d, Symmetriebedingung), muss die Organisation Räume schaffen, in denen diese Entscheidung ohne institutionellen Kostendruck getroffen werden kann. Zu prüfen ist:

- Ist es zulässig, Systemempfehlungen nicht zu folgen? Ist es möglich, ein System in bestimmten Situationen gar nicht zu verwenden?
- Müssen Abweichungen besonders begründet oder dokumentiert werden? Werden Nichtnutzungen statistisch überwacht oder implizit als Qualitätsmangel interpretiert?
- Ist klar geregelt, in welchen Kontexten Nichtnutzung nicht nur zulässig, sondern erforderlich ist?

Diese Regelung muss institutionell verankert sein und nicht etwa lediglich formal erklärt werden.

dd. Qualifikation und gemeinsame Lernprozesse

Der verantwortbare KI-Einsatz setzt ausreichende Kompetenzen bei allen beteiligten Akteursgruppen voraus. Kompetenz entscheidet dabei, ob die epistemische Zugangsbedingung (C.I.4.d.bb) im Alltag eingelöst wird. Zu prüfen ist:

- Werden regelmäßige Schulungen zur technischen Funktionsweise, zu Grenzen und Risiken des Systems angeboten – einschließlich möglicher Bias-Strukturen und Fehlklassifikationen?
- Existieren interdisziplinäre Trainings, in denen fachliche Anwender und technische Entwickler gemeinsam Fallbeispiele analysieren?
- Werden reale Fehlentscheidungen systematisch aufgearbeitet?
- Sind Automation Bias und kognitive Verzerrungen Teil der Ausbildung?

Kompetenzaufbau muss als kontinuierlicher Lernprozess verstanden werden, nicht als einmalige Schulungsmaßnahme. Systeme verändern sich, Nutzungspraktiken entwickeln sich, und neue Risikokonfigurationen entstehen erst im laufenden Betrieb.

ee. Monitoring, Fehlerkultur und kognitive Sicherungsmechanismen

Organisationale Verantwortung erschöpft sich nicht in der Einführungsphase. Sie erfordert Mechanismen, die eine kontinuierliche Evaluation ermöglichen und sicherstellen, dass Fehler systemisch untersucht werden. Zu prüfen ist:

- Wird erfasst, wann Systemempfehlungen übernommen, übersteuert oder nicht genutzt werden? Werden Entscheidungsabweichungen systematisch analysiert?
- Existieren strukturierte Verfahren zur Fehleranalyse (z.B. Root-Cause-Analysen)? Werden Fehler primär einzelnen Personen zugeschrieben oder systemisch untersucht?
- Sind Eskalationsmechanismen und externe Audits vorgesehen?

Ein zentraler Bestandteil organisationaler Sicherungsmechanismen sind Cognitive Forcing Mechanisms – institutionelle Verfahren, die eine aktive Prüfung von Systemempfehlungen fördern, statt sie dem individuellen Ermessen zu überlassen. Zu prüfen ist:

- Müssen Entscheidungen aktiv begründet werden, wenn sie auf einer Systemempfehlung beruhen?
- Werden Alternativhypothesen oder Gegenprüfungen systematisch angeregt?
- Gibt es verpflichtende Plausibilitätsprüfungen bei besonders risikoreichen Entscheidungen?

Solche Mechanismen sind die organisationale Antwort auf das Grundproblem, dass menschliche Kontrolle auch eine institutionelle Leistung ist, die aktiv hergestellt und aufrechterhalten werden muss.

h. Synthese: Zusammenführung der drei Perspektiven

Jede der vorgestellten Perspektiven adressiert eine andere Schicht der Risikogenese. Die Anwenderperspektive bestimmt, welche Risiken aus der Struktur der Entscheidungssituation plausibel sind – abhängig von Normativität, Komplexität und Algorithmisierbarkeit. Die Systemdesigner-Perspektive bestimmt, ob technische

Gestaltungsentscheidungen die in C.I.4.d identifizierten Aufsichtsbedingungen strukturell erfüllen oder unterlaufen. Die organisationale Perspektive bestimmt, ob die institutionellen Rahmenbedingungen wirksame Kontrolle im Alltag ermöglichen oder systematisch erschweren.
Die vier Szenarien in C.I.4.e illustrieren, dass Risiken typischerweise aus dem Zusammenspiel aller drei Ebenen entstehen. Automation Bias (Szenario 1) ist ein kognitives Phänomen auf der Anwenderebene, das durch Interface-Design auf der Designerebene begünstigt und durch fehlende Schulungsstrukturen auf der Organisationsebene verstetigt wird. Institutionelle Standardisierung (Szenario 3) ist ein systemischer Effekt, der durch Rankings, Scores und Default-Optionen auf der Designerebene angestoßen, durch organisationale Effizienzvorgaben verstärkt und erst sichtbar wird, wenn die Anwenderebene bereits angepasstes Verhalten zeigt.
Daraus folgt: Eine Risikobewertung, die nur eine Perspektive einnimmt, erfasst das Risikoprofil einer Anwendung nur unvollständig. Die Kontextanalyse (C.I.4.c) liefert die notwendige Grundlage, ist aber nicht hinreichend, wenn technische Gestaltung oder institutionelle Rahmenbedingungen die identifizierten Risiken verstärken, statt sie einzugrenzen.

aa. Konsequenzen für die Risikobewertung

Für die Bewertung konkreter Anwendungen ergeben sich drei Konsequenzen:
Erstens: Risikobewertung muss kontextspezifisch erfolgen. Dieselbe Technologie kann in einem administrativen Kontext mit geringer Normativität und hoher Formalisierbarkeit ein niedriges Risikoprofil aufweisen und in einem normativen Kontext mit hoher Eingriffsintensität und Einzelfallsensibilität ein hohes. Das Risiko liegt nicht im System als solchem, sondern in der Konfiguration aus System, Entscheidungssituation und institutioneller Einbettung.
Zweitens: Die Wirksamkeit menschlicher Kontrolle ist eine empirisch zu prüfende Bedingung. Die vier Aufsichtsbedingungen (C.I.4.d.bb) müssen für jede konkrete Anwendung geprüft werden – unter Berücksichtigung der tatsächlichen Arbeitsbedingungen, nicht nur der formalen Verfahrensvorgaben. Regulatorische Modelle, die Human Oversight als Sicherheitsmerkmal voraussetzen, ohne dessen Wirksamkeit zu prüfen, greifen zu kurz.
Drittens: Der Evaluationsrahmen ist iterativ anzuwenden. Systeme verändern sich, Nutzungspraktiken entwickeln sich, und neue Risikokonfigurationen entstehen erst im laufenden Betrieb. Die in C.I.4.g.ee beschriebenen Monitoring- und Feedbackstrukturen sind daher nicht optional, sondern konstitutiver Bestandteil der Risikobewertung. Sie liefern zugleich die Daten, die eine quantitative Risikoanalyse im Sinne von C.I.3.b erst ermöglichen.

bb. Offene Anschlüsse

Der Evaluationsrahmen hat die Risikodimensionen, Aufsichtsbedingungen und Risikopfade identifiziert, die für eine kontextspezifische Bewertung des KI-Einsatzes in der richterlichen Entscheidungsfindung maßgeblich sind. Drei Fragen, die sich aus der Analyse ergeben, werden in den folgenden Kapiteln aufgegriffen:

1. Wie lässt sich die Zurechnung von Verantwortung gestalten, wenn Entscheidungen aus dem Zusammenspiel menschlicher und algorithmischer Faktoren entstehen?

2. Unter welchen Bedingungen sind Trade-offs zwischen Effizienzgewinnen und Risikosteigerungen normativ akzeptabel – und wer entscheidet darüber?
3. Welche institutionellen, technischen und verfahrensbezogenen Maßnahmen können die identifizierten Risikopfade wirksam unterbrechen?

Diese Fragen betreffen die normative und institutionelle Konsequenz der Risikoanalyse.

i. Durchlaufendes Anwendungsbeispiel: KI-gestützte Strafzumessungsempfehlung

aa. Beschreibung des Anwendungsbeispiels

Das exemplarische System ist eine KI-gestützte Strafzumessungsempfehlung. Auf Grundlage historischer Falldaten, gesetzlicher Strafrahmen und fallbezogener Merkmale (Deliktart, Vorstrafen, Täter-Opfer-Beziehung, Umstände der Tat) generiert das System eine Empfehlung zur Strafhöhe. Die Empfehlung wird als numerischer Vorschlag mit einem Konfidenzintervall dargestellt. Die formale Letztentscheidung verbleibt beim Gericht.
Dieses Beispiel ist bewusst gewählt, weil es alle drei Risikodimensionen in ausgeprägter Form berührt und die vier typischen Risikopfade (C.I.4.d.dd) unmittelbar einschlägig sind. Es handelt sich um ein fiktives Beispiel, das nicht zu den im ersten Bericht dieses Projekts untersuchten Einsatzszenarien gehört.[44] Es ist gleichwohl ein durchaus denkbares Szenario für den Einsatz von KI in der Justiz: Algorithmische Risikobewertungen fließen in anderen Rechtsordnungen bereits in Strafzumessungsentscheidungen ein (vgl. C.I.5.c.bb zu State v. Loomis). Das Beispiel dient der Illustration des Evaluationsrahmens, nicht der Bewertung eines konkreten Produkts.

[44] Vgl. *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 33ff.

bb. Phase 1: Kontexteinordnung – Anwenderperspektive

Die Anwendung der Reflexionsfragen aus C.I.4.c.aa–C.I.4.c.gg auf das Beispielsystem ergibt folgende Einordnung:

Risikodimensionen

Dimension	Einordnung	Begründung
Funktionaler Charakter	**Hoch normativ**	Betrifft subjektive Rechte (Freiheit), hohe Eingriffsintensität, asymmetrische Machtposition, personengebundene Verantwortung (Art. 92 GG, Art. 97 GG).
Automatisierungsgrad	**Empfehlend bis vorstrukturierend**	Formal empfehlend, aber durch numerische Präzision und Default-Wirkung faktisch vorstrukturierend. Abweichung erfordert Begründungsaufwand.
Komplexität	**Hoch**	Strafzumessung erfordert normative Abwägung, Einzelfallsensibilität, Berücksichtigung situativer Umstände. Bewertungsmaßstäbe sind teilweise offen.

(1) Kontextfaktoren im Detail

Reversibilität (C.I.4.c.bb): Entscheidungsfolgen sind faktisch irreversibel. Freiheitsentzug, soziale Stigmatisierung und wirtschaftliche Konsequenzen lassen sich durch spätere Korrektur nicht vollständig aufheben. Selbst eine erfolgreiche Revision beseitigt nicht die Belastungen der Zwischenzeit.
Erforderliche Fähigkeiten (C.I.4.c.cc): Strafzumessung erfordert normative Abwägung, Empathie, situatives Urteil und Verantwortungsübernahme. Diese Fähigkeiten sind strukturell an menschliche Entscheidungsträger gebunden. Normative Begründung und Rechtfertigung sind integraler Bestandteil der Entscheidung.
Entscheidungsstruktur (C.I.4.c.dd): Strafzumessung beruht wesentlich auf Erfahrungswissen und impliziten Abwägungsprozessen. Entscheidungsregeln sind nur begrenzt explizierbar und konsistent anwendbar. Der Ermessensspielraum ist groß, Grundrechte unmittelbar betroffen.
Entscheidungsvarianz (C.I.4.c.ee): Unterschiede in der Strafzumessung zwischen Gerichten sind für Deutschland empirisch gut dokumentiert. Grundies[45] zeigt auf Grundlage von rund 1,5 Millionen Entscheidungen aus dem Bundeszentralregister deutliche regionale Unterschiede im Strafniveau zwischen Gerichtsbezirken bei vergleichbaren Delikten. Hinzu kommen Ankereffekte, die numerische Entscheidungen auch bei erfahrenen Richterinnen und Richtern systematisch

[45] *Grundies*, Regionale Unterschiede in der gerichtlichen Sanktionspraxis in der Bundesrepublik Deutschland. Eine empirische Analyse, in: Hermann/Pöge (Hrsg.), Kriminalsoziologie – Handbuch für Wissenschaft und Praxis, 1. Aufl., Baden-Baden 2018, S. 295ff.

verschieben können.[46] Ob diese Varianz Ausdruck verzerrter Entscheidungsroutinen oder sachlich gebotener Differenzierung ist, lässt sich nicht pauschal beantworten.
Algorithmisierbarkeit (C.I.4.c.ff): Teilaufgaben wie der Abgleich mit gesetzlichen Strafrahmen oder die Identifikation vergleichbarer Fälle sind formalisierbar. Die zentrale Abwägungsentscheidung – die Gewichtung mildernder und erschwerender Umstände im Einzelfall – entzieht sich jedoch einer zuverlässigen Operationalisierung.
Status quo (C.I.4.b): Auch die rein menschliche Strafzumessung weist die in C.I.4.d.dd beschriebenen Risikopfade auf, nur in anderer Form. Regionale Unterschiede im Strafniveau[47] entsprechen einer unkontrollierten Varianz, Ankereffekte[48], etwa durch den Strafantrag der Staatsanwaltschaft, einer epistemischen Überbewertung einzelner Zahlenwerte. Die Korrekturmöglichkeiten sind begrenzt: Die Strafzumessung ist grundsätzlich Sache des Tatgerichts und revisionsrechtlich nur auf Rechtsfehler überprüfbar. Ein KI-System ist daher nicht an einer fehlerfreien menschlichen Praxis zu messen, sondern daran, ob es diese bekannten Schwächen verringert, ohne neue zu erzeugen. Belastbare Evidenz, die einen solchen Vergleich für den deutschen Strafprozess erlauben würde, liegt allerdings nicht vor.

(2) Risikoprofil

Die Kombination aus hoher Normativität, faktischer Vorstrukturierung und hoher Komplexität ergibt ein **hohes Risikoprofil**. Alle vier typischen Risikopfade (C.I.4.d.dd) sind im konkreten Fall plausibel:

- **Epistemische Überbewertung**: Die numerische Präzision einer Strafempfehlung suggeriert Objektivität, wo tatsächlich Modellunsicherheit besteht.
- **Eingeschränkte Intervention**: Abweichung von einer bezifferten Empfehlung erfordert zusätzlichen Begründungsaufwand.
- **Institutionelle Standardisierung**: Bei breiter Implementierung könnte das System die natürliche Varianz richterlicher Strafzumessung reduzieren, auch dort, wo Einzelfallsensibilität geboten ist.

Fehlinterpretation: Statistische Muster aus historischen Falldaten können als normative Richtwerte missverstanden werden.

(3) Zwischenfazit Anwenderperspektive

Das Risikoprofil spricht gegen eine uneingeschränkte Nutzung als Strafzumessungsempfehlung. Die zentrale Abwägungsentscheidung ist strukturell an menschliche Urteilskraft gebunden und nur begrenzt algorithmisierbar. Ein System, das dennoch in diesem Kontext eingesetzt werden soll, muss besonders strenge Anforderungen an die Aufsichtsbedingungen (C.I.4.d.bb) erfüllen. Zugleich sind

[46] *Bystranowski/Janik/Próchnicki/Skórska*, Anchoring Effect in Legal Decision-Making: A Meta-Analysis, Law and Human Behavior 45 (2021), 1.
[47] *Grundies*, Regionale Unterschiede in der gerichtlichen Sanktionspraxis in der Bundesrepublik Deutschland. Eine empirische Analyse, in: Hermann/Pöge (Hrsg.), Kriminalsoziologie – Handbuch für Wissenschaft und Praxis, 1. Aufl., Baden-Baden 2018, S. 295ff.
[48] *Bystranowski/Janik/Próchnicki/Skórska*, Anchoring Effect in Legal Decision-Making: A Meta-Analysis, Law and Human Behavior 45 (2021), 1.

Teilfunktionen – etwa die Identifikation vergleichbarer Fälle oder der Abgleich mit Strafrahmen – denkbare Einsatzfelder mit deutlich niedrigerem Risikoprofil.

cc. Phase 2: Technische Gestaltung – Designerperspektive

Auf Grundlage des hohen Risikoprofils stellen sich für die Systemgestaltung besonders scharfe Anforderungen. Die folgenden Abschnitte zeigen exemplarisch, welche konkreten Risiken sich auf der Designebene manifestieren und welche Gestaltungsentscheidungen ihnen entgegenwirken können.

(1) Interface-Gestaltung (C.I.4.f.aa)

Risiko: Das System stellt eine Strafempfehlung als numerischen Wert dar (z. B. „2 Jahre, 4 Monate“). Selbst mit Konfidenzintervall erzeugt diese Darstellungsform den Eindruck einer objektiven Berechnung und aktiviert den Risikopfad der epistemischen Überbewertung (Szenario 1).

Abhilfe

- Darstellung als Bandbreite statt als Punktschätzung: „Das Modell identifiziert für vergleichbare Fälle eine Bandbreite von 1 Jahr 8 Monaten bis 3 Jahren.“
- Explizite Unsicherheitskommunikation im Interface: Angabe der Modellgüte, der Größe der Referenzklasse und der Streuung innerhalb vergleichbarer Fälle.
- Keine Darstellung als empfohlene Strafe, sondern als statistische Einordnung: „In 80 % vergleichbarer Fälle lag die Strafe in diesem Bereich.“ Die sprachliche Rahmung bestimmt die epistemische Wirkung.

(2) Verhaltenslenkung (C.I.4.f.bb)

Risiko: Wird die Systemempfehlung als vorausgefüllter Default in der Entscheidungsmaske angezeigt, entsteht eine asymmetrische Entscheidungsarchitektur: Übernahme erfordert einen Klick, Abweichung erfordert Eingabe und Begründung. Dies begünstigt den Risikopfad der eingeschränkten Intervention (Szenario 2).

Abhilfe

- Kein vorausgefüllter Default: Das Strafmaß-Feld bleibt leer. Die Systemeinordnung wird separat angezeigt, nicht als Eingabewert.
- Symmetrischer Begründungsaufwand: Wenn Abweichungen begründet werden müssen, muss auch die Übernahme einer Systemempfehlung kurz begründet werden. Damit wird die Begründungslast neutralisiert.
- Aktive Bestätigung: Vor Übernahme einer Empfehlung muss eine Plausibilitätsprüfung (z. B. Gegenprüfung anhand einer selbst formulierten Einschätzung) dokumentiert werden.

(3) Delegationsgrad (C.I.4.f.cc)

Risiko: Das System generiert einen konkreten Entscheidungsvorschlag. Im Kontext der Strafzumessung überschreitet dies die Grenze zwischen Information und

Empfehlung: Das System nimmt implizit eine Gewichtung mildernder und erschwerender Umstände vor, die den Kern richterlicher Abwägung bildet.

Abhilfe

- Funktionale Beschränkung: Das System liefert keine Strafempfehlung, sondern eine Fallvergleichsanalyse: ähnliche Fälle, deren Ausgänge und die für die Ähnlichkeit maßgeblichen Merkmale. Die Abwägung verbleibt vollständig beim Richter.
- Mindest-Benchmark als Ausschlusskriterium: Wenn das System dennoch als Empfehlungssystem eingesetzt werden soll, muss es nachweislich die Qualität menschlicher Entscheidungen nicht verschlechtern. Mindestanforderungen könnten umfassen: keine sachlich nicht gerechtfertigte Ungleichbehandlung anhand geschützter Merkmale (Fairness-Audit mit offengelegten Kriterien), dokumentierte Modellgüte auf einem repräsentativen Testdatensatz sowie regelmäßige Rekalibrierung bei veränderter Rechtslage.

(4) Eingriffsmöglichkeiten (C.I.4.f.dd)

Risiko: Override ist technisch möglich, erfordert aber zusätzliche Dokumentation. Unter Zeitdruck führt dies dazu, dass Abweichung vermieden wird, selbst wenn der Einzelfall sie erfordert.

Abhilfe

- Override ohne Mehraufwand: Abweichung von der Systemeinordnung darf technisch nicht aufwändiger sein als deren Übernahme.
- Frühe Interventionspunkte: Eingriffsmöglichkeiten vor der Empfehlung, nicht erst danach – etwa durch die Möglichkeit, Fallmerkmale zu korrigieren oder Gewichtungen anzupassen, bevor das System eine Einordnung generiert.

(5) Epistemische Transparenz und Risikokommunikation (C.I.4.f.ee–C.I.4.f.ff)

Risiko: Anwender verstehen nicht, auf welcher Datenbasis und mit welchen Modellannahmen die Empfehlung zustande kommt. Probabilistische Aussagen werden als deterministische Prognosen interpretiert (Szenario 4).

Abhilfe

- Fallbezogene Erklärbarkeit: Für jede Empfehlung muss nachvollziehbar sein, welche Merkmale die Einordnung maßgeblich beeinflusst haben.
- Dokumentierte Modellgrenzen: Klare Angabe, für welche Delikttypen das Modell trainiert wurde und wo die Referenzklasse zu klein für belastbare Aussagen ist.
- Gegenanzeigen: Explizite Benennung von Situationen, in denen das System nicht eingesetzt werden sollte – etwa bei atypischen Fallkonstellationen, die im Trainingsdatensatz unterrepräsentiert sind.

(6) Zwischenfazit Designerperspektive

Die Analyse zeigt, dass ein empfehlendes System zur Strafzumessung nur dann vertretbar ist, wenn die Gestaltung aktiv gegen alle vier Risikopfade absichert. Die wirksamste einzelne Maßnahme ist die funktionale Beschränkung: Reduktion des Systems von einer Strafempfehlung auf eine Fallvergleichsanalyse. Damit wird der Automatisierungsgrad gesenkt, ohne die informativen Teilfunktionen zu verlieren. Wird dennoch an einer Empfehlungsfunktion festgehalten, müssen Mindest-Benchmarks als Ausschlusskriterien definiert werden, bei deren Nichterfüllung das System nicht eingesetzt werden darf.

dd. Phase 3: Institutionelle Einbettung – Organisationale Perspektive

Selbst ein sorgfältig gestaltetes System kann problematische Effekte entfalten, wenn die institutionellen Rahmenbedingungen unzureichend sind. Die organisationale Perspektive prüft, ob die in Phase 2 adressierten technischen Kontrollmöglichkeiten im Alltag wirksam werden.

(1) Erwartungs- und Rollenklarheit (C.I.4.g.aa)

Risiko: Die Einführung des Systems wird von der Justizverwaltung als Effizienzmaßnahme kommuniziert. Es entsteht die implizite Erwartung, dass das System regelmäßig genutzt und seine Empfehlungen in der Regel übernommen werden. Richter, die das System nicht nutzen, werden als ineffizient wahrgenommen. Dies begünstigt den Risikopfad der institutionellen Standardisierung (Szenario 3).

Abhilfe

- Explizite Rollenklarstellung: Schriftliche Festlegung, dass das System ausschließlich informierenden Charakter hat und keine institutionelle Erwartung an die Übernahme seiner Empfehlungen besteht.
- Keine Nutzungsquoten: Nutzungshäufigkeit darf weder erhoben noch als Leistungsindikator verwendet werden.
- Ausschlussbereiche: Klare Definition, in welchen Verfahren das System nicht eingesetzt werden darf – etwa bei Kapitaldelikten, Jugendstrafrecht oder Verfahren mit hohem öffentlichem Interesse.

(2) Kommunikationsstrukturen (C.I.4.g.bb)

Risiko: Richter bemerken systematische Auffälligkeiten in den Empfehlungen (z. B. konsistent höhere Strafen für bestimmte Deliktgruppen), können diese aber nicht strukturiert rückmelden. Informationsasymmetrien zwischen Entwicklern und Anwendern bleiben bestehen.

Abhilfe

- Strukturierter Feedbackkanal: Formalisiertes Verfahren zur Meldung von Auffälligkeiten, mit garantierter Rückmeldung durch das Entwicklerteam.
- Transparente Update-Kommunikation: Jede Änderung an Modell, Datenbasis oder Parametrisierung wird vorab kommuniziert, einschließlich möglicher Auswirkungen auf Empfehlungen.

- Interdisziplinäre Schnittstelle: Regelmäßiger Austausch zwischen Richtern, Justizverwaltung und Systementwicklern – nicht nur bei Einführung, sondern fortlaufend.

(3) Umgang mit Nichtnutzung (C.I.4.g.cc)

Risiko: Ein Richter entscheidet sich im Einzelfall bewusst gegen die Nutzung des Systems, etwa weil die Fallkonstellation atypisch ist. Dies wird in der Fallstatistik als Abweichung sichtbar und zieht Nachfragen der Gerichtsverwaltung nach sich.

Abhilfe

- Institutionell geschützte Nichtnutzung: Nichtnutzung ist kein erklärungsbedürftiger Ausnahmefall, sondern eine jederzeit zulässige Entscheidung.
- Keine statistische Überwachung individueller Nutzungsmuster: Aggregierte Nutzungsdaten können erhoben werden, individuelle Nutzungsprofile nicht.
- Definierte Gegenanzeigen: In Ergänzung zu den technischen Gegenanzeigen (Phase 2) gibt es organisational definierte Situationen, in denen Nichtnutzung nicht nur zulässig, sondern gefordert ist.

(4) Qualifikation (C.I.4.g.dd)

Risiko: Richter erhalten eine einmalige Schulung bei Einführung des Systems, in der die Grundfunktionen erklärt werden. Grenzen des Modells, Automation Bias und die Unterscheidung zwischen statistischer Einordnung und normativer Bewertung werden nicht thematisiert.

Abhilfe

- Schulung mit Fokus auf Systemgrenzen: Neben der Funktionsweise müssen systematisch behandelt werden: die Differenz zwischen Gruppen- und Individualprognose, bekannte Bias-Risiken, Automation Bias und der epistemische Status statistischer Empfehlungen.
- Praxisbezogene Fallbeispiele: Gemeinsame Analyse von Fällen, in denen die Systemempfehlung korrekt war, und Fällen, in denen sie zu einer Fehleinschätzung geführt hätte.
- Regelmäßige Auffrischung: Jede Modelländerung erfordert eine aktualisierte Schulung.

(5) Monitoring und kognitive Sicherungsmechanismen (C.I.4.g.ee)

Risiko: Die Übernahmequote der Systemempfehlungen steigt im Laufe der Zeit von 60 % auf 85%, ohne dass dies als Indikator für nachlassende kritische Prüfung erkannt wird. Gleichzeitig sinkt die Varianz der Strafhöhen – ein möglicher Hinweis auf schleichende Standardisierung.

Abhilfe

- Monitoring-Kennzahlen: Systematische Erhebung der Übernahmequote, der Override-Rate, der Varianz der Entscheidungsergebnisse und der Begründungsdichte bei Übernahme und Abweichung.
- Schwellenwerte als Frühwarnsystem: Definierte Schwellenwerte, bei deren Überschreitung eine institutionelle Prüfung ausgelöst wird – z. B. wenn die Übernahmequote in einem Spruchkörper dauerhaft über 90% liegt (zu den Grenzen solcher Kennzahlen vgl. C.I.5.c.cc).
- Cognitive Forcing Mechanisms: Verpflichtende Eigeneinschätzung vor Einsicht in die Systemempfehlung. Der Richter formuliert zunächst eine eigene vorläufige Bewertung; erst danach wird die Systemeinordnung sichtbar. Damit wird der Ankereffekt (Szenario 1) strukturell unterbrochen.
- Externe Audits: Regelmäßige unabhängige Überprüfung der Systemleistung, der Nutzungsmuster und der Entscheidungsqualität.

(6) Zwischenfazit Organisationsperspektive

Die organisationale Analyse verschärft das Bild: Selbst ein gut gestaltetes System kann unter ungünstigen institutionellen Bedingungen alle vier Risikopfade aktivieren. Die wirksamsten organisationalen Maßnahmen sind: eine explizite Rollenklarstellung ohne Nutzungsquoten, institutionell geschützte Nichtnutzung, Eigeneinschätzung vor Systemeinsicht und Monitoring mit definierten Frühwarnschwellen.

ee. Gesamtfazit: Bewertung des Anwendungsbeispiels

Die Durcharbeitung aller drei Evaluationsebenen ergibt für das exemplarische System ein differenziertes Bild, das sich in einer Stufenlogik zusammenfassen lässt:

Systemfunktion	Risikobewertung	Bedingungen / Maßnahmen
Fallvergleichsanalyse (informierend)	Moderates Risiko. Unterstützt Information, ohne Abwägung vorzustrukturieren.	Epistemische Transparenz, Schulung, Monitoring. Einsatz grundsätzlich vertretbar.
Strafempfehlung mit Bandbreite (empfehlend)	Hohes Risiko. Alle vier Risikopfade plausibel. Automation Bias und institutionelle Standardisierung wahrscheinlich.	Nur unter strengen Bedingungen: Mindest-Benchmarks (Fairness-Audit, Modellgüte), symmetrischer Begründungsaufwand, Eigeneinschätzung vor Systemeinsicht, Monitoring mit Schwellenwerten. Ausschlussbereiche definieren.
Strafempfehlung als Default-Wert (vorstrukturierend)	Sehr hohes Risiko. Faktische Automatisierung bei formaler Letztentscheidung. Meaningful Human Intervention strukturell gefährdet.	Nicht empfohlen. Die Kontextanalyse zeigt, dass diese Konfiguration die Aufsichtsbedingungen (C.I.4.d.bb) systematisch untergräbt. Nichtnutzung ist in diesem Fall die risikoadäquate Entscheidung.

j. Durchlaufendes Anwendungsbeispiel: MAKI (Massenverfahrens-Assistenz mithilfe von KI)

aa. Beschreibung des Anwendungsbeispiels

MAKI (Massenverfahrens-Assistenz mithilfe von KI) ist ein modular aufgebautes KI-Assistenzsystem, das das Niedersächsische Justizministerium seit 2023 entwickelt und das seit Ende 2024 an mehreren niedersächsischen Gerichten von ausgewählten Richterinnen und Richtern erprobt wird.[49] Mit 2024 eingeworbenen Mitteln der Digitalisierungsinitiative des Bundes wird es als MAKI 2.0 weiterentwickelt; Brandenburg, Hessen und Nordrhein-Westfalen haben sich dem Vorhaben 2025 angeschlossen. Die erste Version (MAKI 1.0) beruhte auf dem sogenannten One-Shot-Annotation-Verfahren: Richterinnen und Richter kennzeichneten in einem Musterfall die entscheidungsrelevanten Informationen und die Gründe ihrer Entscheidung; das System erkannte vergleichbare Konstellationen in ähnlichen Fällen und schlug die annotierte Entscheidung dort als anpassbares Muster vor. Daraus sind verschiedene Anwendungen hervorgegangen: in Fluggastrechteverfahren Metadatenextraktion, aktenübergreifender Textvergleich und Vorschläge passender Entscheidungsmuster; in verwaltungsgerichtlichen Asylverfahren der Tatbestandsassistent TabeA, der auf der Extraktionsfunktion von MAKI aufbaut, Kerndaten standardisiert aufbereitet, die vorgetragenen Ausreisegründe zusammenfasst und das Anhörungsprotokoll des BAMF zu Formulierungsvorschlägen für den Tatbestand verdichtet, die vom erkennenden Spruchkörper zu prüfen und gegebenenfalls anzupassen sind.[50] MAKI 2.0 beruht auf großen Sprachmodellen. Eine Kernkomponente bindet das Sprachmodell an und steuert die Abläufe; darauf setzen fachliche Bausteine auf: eine Textbausteinkomponente, die von Richterinnen und Richtern angelegte Textbausteine über Platzhalter mit den entscheidungserheblichen Tatsachen des einzelnen Verfahrens befüllt, eine Strukturierungskomponente für Zusammenfassungen, Relationstabellen und Beweismittelübersichten, eine Dashboardkomponente, die gleichgelagerte Verfahren eines Dezernats bündelt, sowie der „MAKItplace“, über den Gerichte und Länder Vorlagen und Textbausteine teilen können. Entscheidungsvorschläge sieht MAKI 2.0 derzeit nicht vor. TabeA unterscheidet zwischen Schlüsseldaten wie Herkunftsland oder Zustellungsdatum, die unverändert aus der Akte übernommen und mit ihrer Fundstelle verknüpft werden, und Schlüsselinformationen wie dem Vortrag zu den Verfolgungsgründen, die die Richterinnen und Richter selbst markieren und die das System lediglich sprachlich in die passende Form überführt; diese lassen sich nicht unmittelbar mit ihrer Fundstelle verknüpfen. Soweit die folgende Analyse Annotationen und Entscheidungsmuster betrifft, bezieht sie sich auf MAKI 1.0. Die Weiterentwicklung greift einige der dort herausgearbeiteten Risiken bereits auf, etwa durch den Verzicht auf Entscheidungsvorschläge und die richterliche Auswahl der Schlüsselinformationen in TabeA.

---

[49] *Niedersächsisches Justizministerium*, Künstliche Intelligenz und Automation in der Justiz, abrufbar unter https://www.mj.niedersachsen.de/startseite/kunstliche_intelligenz_und_automation_in_der_justiz/kunstliche-intelligenz-und-automation-in-der-justiz-246705.html (Abruf am 04.11.2026, 19:41 Uhr).

[50] *Irskens*, KI in der niedersächsischen Justiz, Anwaltspraxis Magazin vom 09.02.2025, abrufbar unter https://anwaltspraxis-magazin.de/fachbeitraege/elektronischer-rechtsverkehr/2025/02/09/b-ki-in-der-niedersaechsischen-justiz-80/ (Abruf am 04.11.2026, 19:48 Uhr).

Regulatorisch ist diese Modularität von unmittelbarer Bedeutung. KI-Systeme, die Justizbehörden bei der Ermittlung und Auslegung von Sachverhalten und Rechtsvorschriften und bei der Anwendung des Rechts unterstützen sollen, gelten nach Anhang III Nr. 8 lit. a KI-VO grundsätzlich als Hochrisiko-Systeme. Art. 6 Abs. 3 KI-VO nimmt jedoch Systeme aus, die kein erhebliches Risiko bergen, weil sie das Ergebnis der Entscheidungsfindung nicht wesentlich beeinflussen. Das gilt insbesondere, wenn sie lediglich eine eng gefasste Verfahrensaufgabe oder eine vorbereitende Aufgabe erfüllen oder Entscheidungsmuster und Abweichungen davon erkennen, ohne die menschliche Bewertung ohne angemessene Überprüfung zu ersetzen oder zu beeinflussen. Nach Erwägungsgrund 61 sollen zudem rein begleitende Verwaltungstätigkeiten, die die Rechtspflege im Einzelfall nicht beeinträchtigen, nicht als hochriskant gelten. Für MAKI folgt daraus eine differenzierte Einordnung: Metadatenextraktion und aktenübergreifender Textvergleich können als vorbereitende bzw. verfahrensbezogene Aufgaben unter die Ausnahme fallen. Der Vorschlag von Entscheidungsmustern in MAKI 1.0 beeinflusst das Entscheidungsergebnis dagegen typischerweise wesentlich und dürfte die Hochrisiko-Einstufung auslösen; für Tatbestandsentwürfe ist die Einordnung offen, weil der Tatbestand zwar keine rechtliche Würdigung enthält, die Auswahl und Verdichtung des Vortrags diese aber vorprägen kann. Da die Einstufung an das KI-System als Ganzes anknüpft, hängt die regulatorische Behandlung auch davon ab, ob diese Funktionen technisch und organisatorisch als getrennte Systeme betrieben werden. Damit bestätigt bereits die Systematik der KI-Verordnung die zentrale These dieses Abschnitts: Das Risikoprofil ist keine Eigenschaft der Technologie, sondern hängt von Funktion und Einsatzkontext ab.
Das Projekt MAKI eignet sich als durchlaufendes Anwendungsbeispiel, weil es sich um ein reales, öffentlich dokumentiertes System handelt, das an deutschen Gerichten erprobt wird, und zwar in rechtlich und normativ fundamental verschiedenartigen Kontexten. Diese Heterogenität der Einsatzfelder – von standardisierten Fluggastrechteverfahren bis zu existenzsensiblen Asylentscheidungen – macht MAKI zu einem instruktiven Prüffall für eine zentrale These des Evaluationsrahmens: Das Risikoprofil eines KI-Systems ist keine stabile Eigenschaft des Systems selbst, sondern eine Funktion der Entscheidungsumwelt, in der es eingesetzt wird.

bb. Phase 1: Kontexteinordnung – Anwenderperspektive

Die Anwendung der Reflexionsfragen aus C.I.4.c.aa–C.I.4.c.gg auf MAKI führt zu erheblich verschiedenen Einordnungen je nach Einsatzkontext. Zwei repräsentative Kontexte werden gegenübergestellt: Fluggastrechteklagen nach der EU-Fluggastrechteverordnung (VO 261/2004) und verwaltungsgerichtliche Asylverfahren. Für das Asylverfahren betrachtet die folgende Analyse zwei Ausbaustufen: die derzeit erprobte Funktion von TabeA, die Tatbestandsentwürfe aus Akteninhalt und Anhörungsprotokoll erstellt, und eine naheliegende Erweiterung, in der das MAKI-Framework – wie MAKI 1.0 im Fluggastrechteverfahren – auch Entscheidungsmuster vorschlägt. Bereits die erste Stufe ist risikorelevant: Zwar wählen die Richterinnen und Richter die maßgeblichen Passagen selbst aus; die sprachliche Verdichtung der vorgetragenen Ausreisegründe und des Anhörungsprotokolls durch das System prägt aber die spätere Glaubhaftigkeitsprüfung, ohne als Wertung erkennbar zu sein, weil Details, Widersprüche oder Steigerungen im Vortrag in einer Zusammenfassung geglättet oder ausgelassen werden können. Die zweite Stufe würde darüber hinaus die normativ entscheidende Abwägung selbst vorstrukturieren.

Risikodimensionen

Dimension	Fluggastrechteklagen	Begründung	Asylverfahren	Begründung
Funktionaler Charakter	**Niedrig bis mittel normativ**	Betrifft primär wirtschaftliche Rechtsposition; Anspruchsgrundlagen durch Verordnung und EuGH-Rechtsprechung standardisiert, Ausgleichshöhe pauschaliert (Art. 7 VO 261/2004).	**Hoch normativ**	Betrifft existenzielle Güter: Schutz vor Verfolgung und ernsthaftem Schaden (§§ 3, 4 AsylG; Art. 16a GG), Refoulement-Verbot (Art. 33 GFK, Art. 3 EMRK). Hohe Eingriffsintensität, asymmetrische Machtposition.
Automatisierungsgrad	**Informierend bis empfehlend**	Entscheidungsvorschlag auf Basis formaler Kriterien; Abweichung wenig komplex.	**Empfehlend bis vorstrukturierend**	Entscheidungsvorschlag kann normativ entscheidende Kernabwägung vorwegnehmen; unter Verfahrensdruck faktisch vorstrukturierend.
Komplexität	**Niedrig bis mittel**	Tatbestandsmerkmale (Verspätungsdauer, Ausfall, außergewöhnliche Umstände) subsumtionsfähig; Beurteilungsspielraum durch umfangreiche EuGH-Rechtsprechung weitgehend konturiert.	**Hoch**	Ländersituationsberichte, individuelle Verfolgungsgeschichten, Glaubwürdigkeitsbeurteilungen erfordern Einzelfallabwägung; Bewertungsmaßstäbe kontextsensitiv und teilweise offen.

(1) Kontextfaktoren im Detail

Reversibilität (C.I.4.c.bb): Im Fluggastrechteverfahren sind Entscheidungsfolgen begrenzt und rein finanzieller Natur; Rechtsmittel sind wegen geringer Streitwerte allerdings oft nicht eröffnet (§ 511 Abs. 2 ZPO). Im Asylverfahren können Entscheidungsfolgen faktisch irreversibel sein: Eine Abschiebung lässt sich auch bei späterer gerichtlicher Korrektur häufig nicht vollständig rückgängig machen. Das Fehlerpotenzial des Systems ist damit kontextabhängig stark asymmetrisch.

Erforderliche Fähigkeiten (C.I.4.c.cc): Fluggastrechteklagen erfordern überwiegend formale Subsumtion anhand gesetzlich fixierter Kriterien. Asylverfahren erfordern darüber hinaus normative Abwägung, Glaubwürdigkeitsbeurteilung, situatives Urteil und Verantwortungsübernahme. Diese Fähigkeiten sind strukturell an menschliche Entscheidungsträger gebunden; normative Begründung und Rechtfertigung sind integraler Bestandteil der Entscheidung.
Entscheidungsstruktur (C.I.4.c.dd): Im standardisierten Fluggastrechteverfahren sind Entscheidungsregeln weitgehend explizierbar und konsistent anwendbar. Im Asylrecht beruht die normativ entscheidende Bewertung wesentlich auf Erfahrungswissen, Aktenkenntnis und impliziten Abwägungsprozessen; der Spielraum bei Tatsachenwürdigung und Prognose ist groß, Grundrechte sind unmittelbar betroffen.
Entscheidungsvarianz (C.I.4.c.ee): Fluggastrechteklagen weisen eine geringe legitime Entscheidungsvarianz auf; eine durch das System induzierte Standardisierung ist hier bereits durch Verordnungsgeber und EuGH-Rechtsprechung weitestgehend vorgenommen worden. Asylverfahren erfordern hingegen fallindividuelle Differenzierung; eine systeminduzierte Reduktion der Varianz wäre normativ problematisch, da sie gerade diejenigen Unterschiede nivellieren könnte, die sachlich geboten sind.
Algorithmisierbarkeit (C.I.4.c.ff): Teilaufgaben wie Metadatenextraktion, Streckenzuordnung und Berechnung gesetzlicher Entschädigungsbeträge im Fluggastrechteverfahren sind vergleichsweise gut formalisierbar. Die normativ entscheidenden Kernfragen im Asylverfahren wie Glaubwürdigkeit individueller Verfolgungsgeschichten, Beurteilung von Ländersituationsberichten und kontextsensitive Schutzbedarfsbeurteilung entziehen sich einer zuverlässigen Operationalisierung.
Status quo (C.I.4.b): Im Fluggastrechteverfahren wird bereits heute in hohem Maß mit Textbausteinen und Mustervorlagen gearbeitet; die typischen Fehler des Status quo sind Übertragungs- und Zuordnungsfehler bei hoher Fallzahl. Ein System, das Metadaten zuverlässig extrahiert, konkurriert hier nicht mit einer individuellen Einzelfallprüfung, sondern mit einer bereits standardisierten Praxis. Im Asylverfahren beruht der Tatbestand schon heute weitgehend auf dem Akteninhalt und der Darstellung im Bescheid des BAMF; das Asylgesetz erlaubt dem Gericht ausdrücklich, auf die Feststellungen des Bescheids Bezug zu nehmen, soweit es ihnen folgt (§ 77 AsylG). Auch die menschliche Praxis verdichtet den Vortrag also bereits, nur ohne systematische Prüfung der Quelltreue. Ob TabeA die Verdichtung des Vortrags gegenüber diesem Status quo verschlechtert oder – etwa durch die Verknüpfung der Schlüsseldaten mit ihrer Fundstelle – sogar nachvollziehbarer macht, ist eine empirische Frage, die die Erprobung beantworten sollte. Für die Schlüsselinformationen ist eine solche Verknüpfung nicht möglich; hier bleibt die Prüfung der Quelltreue durch den Spruchkörper unverzichtbar.

(2) Risikoprofil

Die Kombination der Dimensionen ergibt nach den in C.I.4.d definierten Grundsätzen für die beiden Kontexte grundlegend verschiedene Risikoprofile.
Im Fluggastrechteverfahren ergibt sich ein **niedriges bis mittleres Risikoprofil**. Die vier typischen Risikopfade sind zwar nicht ausgeschlossen, aber in ihrer Wirkung begrenzt:

- **Epistemische Überbewertung**: Begrenzt durch hohen Formalisierungsgrad; numerische Outputs entsprechen weitgehend der tatsächlichen Aufgabenstruktur.
- **Eingeschränkte Intervention**: Gering, da Abweichung keine komplexe Begründung erfordert und die Entscheidungskriterien eng und überprüfbar sind.
- **Institutionelle Standardisierung**: In Grenzen wünschenswert; Standardisierung dient hier der Rechtsvereinheitlichung.
- **Fehlinterpretation**: Begrenzt, da die Vergleichbarkeit historischer Fälle bei formalisierten Kriterien hoch und die Übertragbarkeit auf neue Fälle gut abschätzbar ist.

Im Asylverfahren ergibt sich ein **hohes bis sehr hohes Risikoprofil**. Alle vier Risikopfade sind plausibel und in ihrer Wirkung erheblich:

- **Epistemische Überbewertung**: Ein sprachlich geschlossener Entscheidungsvorschlag oder eine generierte Tatbestandsformulierung suggeriert Faktizität, wo tatsächlich eine statistische Ähnlichkeitsbeurteilung auf Grundlage vergangener Annotationen vorliegt.
- **Eingeschränkte Intervention**: Unter dem Verfahrensdruck von Massenverfahren kann die Abweichung von einem formulierten Entscheidungsvorschlag einen erheblichen kognitiven und dokumentarischen Mehraufwand bedeuten; die Begründungslast verschiebt sich faktisch.
- **Institutionelle Standardisierung**: Das Annotationsverfahren kodiert die normative Perspektive einzelner Richterinnen und Richter; seine Übernahme durch andere Richterinnen und Richter ohne eigene Prüfung kann individuelle normative Vorentscheidungen institutionell wirksam machen, ohne dass dies explizit entschieden wurde.
- **Fehlinterpretation**: Ländersituationen verändern sich; individuelle Verfolgungsgeschichten sind strukturell singulär. Ein hoher Ähnlichkeitsscore mit historischen Fällen ist in diesem Kontext kein Indikator für normative Richtigkeit im Einzelfall, sondern allenfalls ein Hinweis auf strukturelle Vergleichbarkeit.

(3) Zwischenfazit Anwenderperspektive

Das Risikoprofil ist kein Merkmal des MAKI-Systems, sondern eine Funktion des Entscheidungskontexts, in dem es eingesetzt wird. Im Fluggastrechteverfahren ist eine KI-Unterstützung mit niedrigem bis mittlerem Risikoprofil vereinbar. Im Asylverfahren ergibt dieselbe Systemfunktion ein hohes Risikoprofil, das besonders strenge Anforderungen an Gestaltung und institutionelle Einbettung stellt. Eine pauschale Einordnung von MAKI als risikoarm oder risikoreich verfehlt diese Differenzierung; sie ist die Voraussetzung für jede weitergehende Evaluierung.

cc. Phase 2: Technische Gestaltung – Designerperspektive

Auf Grundlage des kontextspezifischen Risikoprofils stellen sich für die Systemgestaltung kontextdifferenzierte Anforderungen. Die folgenden Abschnitte zeigen exemplarisch, welche konkreten Risiken sich auf der Designebene manifestieren und welche Gestaltungsentscheidungen ihnen entgegenwirken können.

Da das hohe Risikoprofil des Asylverfahrens die strengeren Anforderungen erzeugt, orientiert sich die Analyse vorrangig an diesem Kontext. Soweit die folgenden Risiken an Annotationen und Entscheidungsmuster anknüpfen, betreffen sie MAKI 1.0.

(1) Interface-Gestaltung (C.I.4.f.aa)

Risiko: MAKI kann sprachlich geschlossene Tatbestandsformulierungen und Entscheidungsvorschläge generieren. Diese erscheinen im Interface als fertige Texte, nicht als Ähnlichkeitssignale auf der Grundlage vergangener Annotationen. Die Darstellungsform aktiviert den Risikopfad der epistemischen Überbewertung (Szenario 1): Ein sprachlich vollständiger Vorschlag suggeriert Faktizität, wo tatsächlich eine statistische Mustererkennung vorliegt.

Abhilfe

- Explizite Herkunftskennzeichnung: Jeder Systemoutput muss seine Entstehungslogik transparent machen – etwa: „Dieser Textvorschlag basiert auf X ähnlichen Fällen; die Ähnlichkeit bezieht sich auf folgende Merkmale: […].“
- Offenlegung der Unsicherheitsfaktoren: Statt eines zusammenfassenden Scores werden die Faktoren sichtbar gemacht, die die Unsicherheit eines Vorschlags begründen, insbesondere die Größe der Referenzklasse und die Streuung innerhalb ähnlicher Fälle. Ist die Referenzklasse klein oder heterogen, wird dies explizit als Einschränkung angezeigt.
- Visuelle Trennung: Systemgenerierte Textbausteine und richterliche Formulierungen werden im Interface klar unterschieden, bis der Richter den Text aktiv für den vorliegenden Fall übernimmt.

(2) Verhaltenslenkung (C.I.4.f.bb)

Risiko: Das Annotationsverfahren von MAKI erzeugt einen selbstreferentiellen Rückkopplungsprozess: Das System identifiziert Muster, die dem eigenen früheren Entscheidungsverhalten entsprechen, und stärkt durch seinen Output die Relevanz genau dieser Muster. Gleichzeitig begünstigt die Effizienzlogik der Massenverfahren die Übernahme von Systemvorschlägen: Abweichung erfordert kognitiven Mehraufwand und zusätzliche Begründung, während Übernahme den effizienten Pfad darstellt. Diese asymmetrische Entscheidungsarchitektur aktiviert den Risikopfad der eingeschränkten Intervention (Szenario 2).

Abhilfe

- Kein vorausgefüllter Default: Systemvorschläge werden separat angezeigt, nicht als Eingabewert in die Entscheidungsmaske übernommen.
- Symmetrischer Begründungsaufwand: Wenn Abweichungen dokumentiert werden müssen, muss auch die Übernahme eines Systemvorschlags eine kurze Plausibilitätskennzeichnung erfordern. Die Begründungslast darf nicht asymmetrisch zur Abweichung hin verschoben sein.
- Eigenformulierung vor Systemeinsicht: Im Asylverfahren sollte der Richter eine eigene vorläufige Einschätzung formulieren, bevor der Systemvorschlag sichtbar wird. Damit wird der Ankereffekt strukturell unterbrochen.

(3) Delegationsgrad (C.I.4.f.cc)

Risiko: Die Funktion der Entscheidungsvorschläge und Tatbestandsformulierungen überschreitet die Grenze zwischen Information und Empfehlung: Das System nimmt implizit eine Vorstrukturierung der normativ entscheidenden Abwägung vor, die im Asylverfahren den Kern richterlicher Prüfung bildet. Eine Tatbestandsformulierung, die auf Annotationen vergangener Fälle basiert, ist keine neutrale Informationsaufbereitung, sondern trägt normative Vorentscheidungen in sich.

Abhilfe

- Funktionale Differenzierung nach Kontext: Im formalisierten Fluggastrechteverfahren kann eine Empfehlungsfunktion vertretbar sein. Im Asylverfahren sollte die Funktion auf Informationsaufbereitung und Metadatenextraktion beschränkt werden; normative Abwägungen verbleiben vollständig beim Richter.
- Modulare Zugangsbeschränkung: Entscheidungsvorschläge und Tatbestandsformulierungen sollten im Asylverfahren nur dann zugänglich sein, wenn eine explizite Entscheidung vorliegt, die die Nutzung freischaltet und ihre Bedingungen definiert.
- Mindest-Benchmarks als Ausschlusskriterium: Wo eine Empfehlungsfunktion eingesetzt wird, muss nachweislich gelten: keine sachlich nicht gerechtfertigte Ungleichbehandlung anhand geschützter Merkmale (Fairness-Audit mit offengelegten Kriterien), dokumentierte Modellgüte auf einem repräsentativen Testdatensatz sowie regelmäßige Rekalibrierung bei veränderter Rechtslage und aktualisierter Ländersituation.

(4) Eingriffsmöglichkeiten (C.I.4.f.dd)

Risiko: Override ist technisch möglich, kann aber unter dem Verfahrensdruck der Massenbearbeitung faktisch mit Mehraufwand verbunden sein. Ist eine abweichende Entscheidung dokumentationspflichtig, kehrt sich die faktische Begründungslast um. Im Asylverfahren ist dieses Risiko besonders virulent, weil gerade hier Einzelfallabwägungen normativ geboten sind.

Abhilfe

- Override ohne Mehraufwand: Abweichung von Systemvorschlägen darf technisch nicht aufwändiger sein als deren Übernahme.
- Frühe Interventionspunkte: Eingriffsmöglichkeiten sollten vor der Generierung des Systemvorschlags bestehen – etwa durch die Möglichkeit, die Merkmale des Referenzfalls zu korrigieren oder die Annotationslogik anzupassen, bevor das System eine Einordnung generiert.
- Explizite Gegenanzeigen: Das System zeigt aktiv an, in welchen Fallkonstellationen eine Nutzung des Vorschlags nicht ratsam ist – etwa wenn die Referenzklasse zu klein ist, wenn der Fall strukturell singuläre Merkmale aufweist oder wenn sich die Ländersituation seit der letzten Annotation verändert hat.

(5) Epistemische Transparenz und Risikokommunikation (C.I.4.f.ee–C.I.4.f.ff)

Risiko: Das Anlernverfahren von MAKI basiert auf Mustererkennung: Das System beurteilt, inwieweit ein neuer Fall den Merkmalen ähnelt, die in vergangenen Fällen als entscheidungsrelevant annotiert wurden. Dieser probabilistische Output tritt im Interface jedoch als sprachlich geschlossener Text auf. Anwender können dazu neigen, ihn als inhaltliche Bestätigung zu interpretieren, nicht als Ähnlichkeitssignal auf der Grundlage vergangener Annotationen. Im Asylrecht ist dieses Risiko besonders ausgeprägt, da historische Entscheidungsmuster nur begrenzt auf neue Fälle übertragbar sind (Szenario 4). Der Mechanismus wird durch den selbstreferentiellen Charakter der Annotation verstärkt: Gerade weil die Systemempfehlungen auf den eigenen Annotationen basieren, können sie als besonders verlässlich wahrgenommen werden.

Abhilfe

- Fallbezogene Erklärbarkeit: Für jeden Systemvorschlag muss nachvollziehbar sein, welche Merkmale die Einordnung maßgeblich beeinflusst haben und welche Fälle als Referenz dienten.
- Dokumentierte Modellgrenzen: Klare Angabe, für welche Falltypen das Modell gelernt hat und wo die Referenzklasse zu klein oder zu heterogen für belastbare Aussagen ist.
- Kontextspezifische Gegenanzeigen: Explizite Benennung von Situationen, in denen das System nicht oder nur eingeschränkt genutzt werden sollte – insbesondere bei veränderter Ländersituation, atypischen individuellen Verfolgungsgeschichten oder Konstellationen, die im Annotationsbestand unterrepräsentiert sind.
- MAKI 1.0 realisierte bereits eine Form der Ähnlichkeitstransparenz. Deren konsequente Umsetzung und systematische Kommunikation an die Anwender sind eine notwendige, aber nicht hinreichende Bedingung für risikoarmen Einsatz.

(6) Zwischenfazit Designerperspektive

Die Analyse zeigt, dass die generische MAKI-Architektur – ein System, das sich durch einmaliges Anlernen auf sehr unterschiedliche Kontexte konfigurieren lässt – eine kontextspezifische Differenzierung im Systemdesign erfordert. Die wirksamste einzelne Maßnahme ist die funktionale Differenzierung nach Kontext: Im formalisierten Fluggastrechteverfahren kann eine Empfehlungsfunktion mit strengen Transparenzanforderungen vertretbar sein; im Asylverfahren ist eine Beschränkung auf Informationsaufbereitung und Metadatenextraktion geboten. Eine einheitliche Designentscheidung für alle Einsatzkontexte verfehlt die Differenzierungsanforderung des Risikoprofils.

dd. Phase 3: Institutionelle Einbettung – Organisationale Perspektive

Selbst ein sorgfältig gestaltetes System kann problematische Effekte entfalten, wenn die institutionellen Rahmenbedingungen unzureichend sind. Die organisationale Perspektive prüft, ob die in Phase 2 adressierten technischen Kontrollmöglichkeiten im Alltag wirksam werden.

(1) Erwartungs- und Rollenklarheit (C.I.4.g.aa)

Risiko: MAKI ist als Effizienzwerkzeug für Massenverfahren konzipiert. Diese Ausgangskonstellation erzeugt institutionelle Erwartungen, die eine schnelle Übernahme von Systemvorschlägen begünstigen können. Ist eine abweichende Entscheidung mit zusätzlichem Dokumentationsaufwand verbunden, kehrt sich die faktische Begründungslast um: Übernahme bleibt der effiziente Pfad, Abweichung wird erklärungsbedürftig. Im Asylverfahren begünstigt dieser Mechanismus den Risikopfad der institutionellen Standardisierung (Szenario 3), weil gerade hier die normative Notwendigkeit zur Einzelfallprüfung am größten ist.

Abhilfe

- Explizite Rollenklarstellung: Schriftliche Festlegung, dass MAKI ausschließlich unterstützenden Charakter hat und keine institutionelle Erwartung an die Übernahme von Systemvorschlägen besteht.
- Kontextspezifische Ausschlussbereiche: Klare Definition, in welchen Verfahrensarten und Entscheidungsphasen das System eingesetzt werden darf und in welchen nicht – insbesondere im Asylverfahren für die normativ entscheidenden Abwägungsschritte.
- Keine Nutzungsquoten: Nutzungshäufigkeit darf weder erhoben noch als Leistungsindikator verwendet werden.

(2) Kommunikationsstrukturen (C.I.4.g.bb)

Risiko: Die länderübergreifende Kooperation und der generische Charakter von MAKI erzeugen eine institutionelle Distanz zwischen Entwicklern und Anwendern. Richterinnen und Richter bemerken möglicherweise Auffälligkeiten im Systemverhalten – etwa systematisch verzerrte Ähnlichkeitseinschätzungen bei bestimmten Herkunftsländern oder Fallkonstellationen –, können diese aber nicht strukturiert zurückmelden. Informationsasymmetrien zwischen Entwicklern und Anwendern bleiben bestehen.

Abhilfe

- Strukturierter Feedbackkanal: Formalisiertes Verfahren zur Meldung von Auffälligkeiten mit garantierter Rückmeldung durch das Entwicklungskonsortium.
- Transparente Update-Kommunikation: Jede Änderung an Modell, Datenbasis oder Annotationslogik wird vorab kommuniziert, einschließlich möglicher Auswirkungen auf bestehende Annotationen und Vorschlagsqualität.
- Interdisziplinäre Schnittstelle: Implementierung eines regelmäßigen Austauschs zwischen Richterinnen und Richtern, Justizverwaltung und Entwicklern.
- Wissenschaftliche Fundierung: Berücksichtigung wissenschaftlicher Forschung, nicht nur bis zur Einführung eines Systems, sondern fortlaufend.

(3) Umgang mit Nichtnutzung (C.I.4.g.cc)

Risiko: Ein Richter entscheidet sich im Einzelfall bewusst gegen die Nutzung des Systems – etwa weil die Fallkonstellation atypisch ist, weil sich die Ländersituation verändert hat oder weil die Referenzklasse zu klein erscheint. Ist diese Entscheidung in der Systemnutzungsstatistik als Abweichung sichtbar und zieht Nachfragen der Gerichtsverwaltung nach sich, entsteht ein faktischer Nutzungsdruck, der mit der Idee der freiwilligen und kontrollbewussten Nutzung unvereinbar ist.

Abhilfe

- Institutionell geschützte Nichtnutzung: Nichtnutzung ist kein erklärungsbedürftiger Ausnahmefall, sondern eine jederzeit zulässige Entscheidung.
- Keine statistische Überwachung individueller Nutzungsmuster: Aggregierte Nutzungsdaten können für die Systemweiterentwicklung erhoben werden, individuelle Nutzungsprofile nicht.
- Definierte Gegenanzeigen: In Ergänzung zu den technischen Gegenanzeigen (Phase 2) gibt es organisational definierte Situationen, in denen Nichtnutzung nicht nur zulässig, sondern geboten ist – insbesondere im Asylverfahren bei Fällen mit ungewöhnlichen Verfolgungskontexten oder unzureichender Referenzklasse.

(4) Qualifikation (C.I.4.g.dd)

Risiko: Richterinnen und Richter erhalten eine Einführungsschulung bei Systemstart, in der die Grundfunktionen erklärt werden. Die spezifischen epistemischen Eigenschaften des One-Shot-Annotation-Verfahrens, insbesondere dass der Systemoutput eine Ähnlichkeitsbeurteilung auf der Grundlage vergangener eigener Annotationen ist und keine faktische Aussage über den vorliegenden Fall, werden dabei möglicherweise nicht ausreichend vermittelt. Dies erhöht die Anfälligkeit für epistemische Überbewertung (Szenario 1) und Fehlinterpretation (Szenario 4).

Abhilfe

- Schulung mit Fokus auf Systemlogik und -grenzen: Neben der Funktionsweise müssen systematisch behandelt werden: die Unterscheidung zwischen Ähnlichkeitssignal und inhaltlicher Bestätigung, der selbstreferentielle Charakter individueller Annotationen, bekannte Grenzen des Mustererkennungsverfahrens bei singulärem Fallmaterial sowie Automation Bias.
- Praxisbezogene Fallbeispiele: Gemeinsame Analyse von Fällen, in denen der Systemvorschlag zutreffend war, und Fällen, in denen er zu einer Fehleinschätzung geführt hätte, insbesondere bei veränderten Ländersituationen oder atypischen Verfolgungsgeschichten.
- Regelmäßige Auffrischung: Jede Modelländerung und jede Erweiterung auf neue Falltypen erfordert eine aktualisierte Schulung.

(5) Monitoring und kognitive Sicherungsmechanismen (C.I.4.g.ee)

Risiko: Eine zentrale Governance-Schwäche liegt darin, dass keine systematische Erhebung darüber erfolgt, in wie vielen Verfahren MAKI tatsächlich genutzt wird. Ohne empirische Grundlage lässt sich weder erkennen, ob Annotationspraktiken zwischen Nutzern konvergieren, noch ob Skalierungseffekte der institutionellen Standardisierung bereits entstehen. Gleichzeitig ermöglicht das individuelle Annotationsverfahren, dass idiosynkratische Entscheidungspraktiken systematisch in den Systemoutput eingebettet werden, ohne dass dies nach außen sichtbar wird.
Besondere Aufmerksamkeit verdient die Frage der Annotationsausweitung: Solange Annotationen ausschließlich auf individueller Ebene verbleiben, ist der Übergang vom idiosynkratischen zum systemischen Risiko strukturell begrenzt. Werden Annotationen dagegen zwischen Richterinnen und Richtern geteilt – innerhalb einer Kammer oder zwischen Einzelrichterinnen und Einzelrichtern desselben Gerichts –, entsteht eine risikokritische Schwelle: Übernehmen andere eine fremde Annotation, ohne sie selbst zu prüfen, werden die normativen Vorentscheidungen einer einzelnen Person in fremden Verfahren wirksam. Dies berührt die richterliche Unabhängigkeit, die auch innerhalb eines Gerichts und einer Kammer gilt.[51] Das Risiko besteht nicht in der geteilten Annotation als solcher, sondern in ihrer ungeprüften Übernahme. Es ist gering, wenn jede nutzende Person die Annotation einsehen, prüfen und anpassen kann und dies tatsächlich tut. Diese Governance-Frage ist keine technische, sondern eine institutionelle Weichenstellung. In MAKI 2.0 stellt sich dieselbe Frage beim „MAKItplace“, über den Gerichte und Länder Vorlagen und Textbausteine teilen können. Zu unterscheiden ist dies von der Konzeption sog. Automatisierungskammern für den entscheidungsersetzenden Einsatz, um die es hier nicht geht.[52]

Abhilfe

- Monitoring-Kennzahlen: Systematische Erhebung der Nutzungsrate, der Override-Rate, der Varianz der Entscheidungsergebnisse und der Übereinstimmung von Systemvorschlägen mit richterlichen Entscheidungen; disaggregiert nach Falltyp und Verfahrensart.
- Annotationsmonitoring: Regelmäßige Überprüfung, ob Annotationspraktiken zwischen Nutzern konvergieren und ob Annotationen ohne eigene Prüfung von anderen übernommen werden, ohne dass dies explizit entschieden wurde.
- Schwellenwerte als Frühwarnsystem: Definierte Schwellenwerte, bei deren Überschreitung eine institutionelle Prüfung ausgelöst wird – etwa wenn die Übernahmequote in einem Spruchkörper dauerhaft über einem definierten Schwellenwert liegt oder wenn die Entscheidungsvarianz in einer Verfahrensart signifikant sinkt (zu den Grenzen solcher Kennzahlen vgl. C.I.5.c.cc).
- Cognitive Forcing Mechanisms im Asylverfahren: Verpflichtende Eigeneinschätzung vor Einsicht in den Systemvorschlag. Die Richterin oder der Richter formuliert zunächst eine eigene vorläufige Bewertung; erst danach wird der Systemvorschlag sichtbar. Damit wird der Ankereffekt strukturell unterbrochen.
- Institutionelle Regelung der Annotationsausweitung: Explizite Festlegung, ob, unter welchen Bedingungen und auf wessen Veranlassung eine Ausweitung

[51] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 192.
[52] Vgl. hierzu *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 219ff.

individueller Annotationen auf andere Richter und Richterinnen zulässig ist. Diese Entscheidung muss transparent geführt und dokumentiert werden.

(6) Zwischenfazit Organisationsperspektive

Die organisationale Analyse verschärft das Bild der Designperspektive: Selbst ein sorgfältig gestaltetes System kann unter institutionellen Effizienzerwartungen alle vier Risikopfade aktivieren. Die wirksamsten organisationalen Maßnahmen sind: kontextspezifische Ausschlussbereiche, institutionell geschützte Nichtnutzung, Eigeneinschätzung vor Systemvorschlag im Asylverfahren sowie Monitoring, das Annotationskonvergenz und Standardisierungseffekte sichtbar macht. Besonders zentral ist die Etablierung klarer Governance-Strukturen für die Frage, ob und unter welchen Bedingungen eine Ausweitung individueller Annotationen von anderen Richtern und Richterinnen übernommen werden dürfen.

ee. Gesamtfazit: Bewertung des Anwendungsbeispiels

Die Durcharbeitung aller drei Evaluationsebenen ergibt für MAKI ein differenziertes Bild, das sich nicht in einer einheitlichen, sondern in einer kontextspezifischen Stufenlogik zusammenfassen lässt:

Systemfunktion	Kontext	Risikobewertung	Bedingungen / Maßnahmen
Metadatenextraktion, Textvergleich (informierend)	Alle Einsatzkontexte	Niedriges bis moderates Risiko. Unterstützt Informationsgewinnung, ohne normative Abwägung vorzustrukturieren.	Epistemische Transparenz zur Herkunft der Extraktion. Kontrolle der korrekten Zuordnung von Schriftsätzen. Einsatz grundsätzlich vertretbar.
Entscheidungsvorschläge (MAKI 1.0), Textbausteine (empfehlend)	Fluggastrechteklagen	Moderates Risiko. Risikopfade begrenzt wirksam bei formalisierten Kriterien; Standardisierung in Grenzen sachgemäß.	Unsicherheitskommunikation, symmetrischer Begründungsaufwand, Monitoring. Einsatz unter diesen Bedingungen vertretbar.
Textbausteinkomponente (MAKI 2.0): Befüllung richterlich angelegter Textbausteine über Platzhalter (informierend bis vorstrukturierend)	Massenverfahren, z. B. Fluggastrechteklagen	Niedriges bis moderates Risiko. Die normative Entscheidung liegt im richterlich angelegten Baustein; Risiken liegen in fehlerhaft eingesetzten Tatsachen und in Transformations-	Kennzeichnung befüllter Stellen bis zur Übernahme, Fundstellenverknüpfung für Schlüsseldaten, gesonderte Prüfung transformierter Inhalte. Einsatz unter diesen Bedingungen vertretbar.

Systemfunktion	Kontext	Risikobewertung	Bedingungen / Maßnahmen
		platzhaltern, deren sprachliche Umformung sich nicht unmittelbar auf die Akte zurückführen lässt (Szenarien 1 und 4).	
Strukturierungs-komponente (MAKI 2.0): Zusammenfas-sungen, Relationstabellen, Beweismittelüber-sichten (informierend bis vorstrukturierend)	Alle Einsatzkontex-te, insbesondere Asylverfahren	Moderates bis hohes Risiko. Zusammenfassun-gen können Inhalte auslassen oder hinzufügen; Relationstabellen ordnen den Vortrag Tatbestandsmerk-malen zu und wirken damit inhaltlich-normativ.	Prüfung der Quelltreue, Kennzeichnung als Systemausgabe, Behandlung von Relationstabellen als inhaltlich-normative Funktion (C.I.4.c.aa), keine systemseitige Auswahl entscheidungsrelevant er Informationen.
Tatbestandsent-würfe	Asylverfahren	Mittleres bis hohes Risiko. Selektive Verdichtung des Vortrags kann die Glaubhaftigkeits-prüfung unbemerkt prägen (Szenario 4).	Nachweis der Quelltreue für Zusammenfassungen, Verknüpfung der Schlüsseldaten mit ihrer Fundstelle, Prüfung der Schlüsselinformatio-nen durch den Spruchkörper, Kennzeichnung als Systementwurf, Schulung.
Entscheidungs-vorschläge (MAKI 1.0) (empfehlend)	Asylverfahren	Hohes Risiko. Alle vier Risikopfade plausibel; normative Vorstrukturierung bei faktischer Irreversibilität der Entscheidungsfol-gen.	Nur unter strengen Bedingungen: Cognitive Forcing, Fairness-Audit, Referenzklassentrans-parenz, Monitoring mit Schwellenwerten, Annotationsmonitoring , explizite Ausschlussbereiche.
Geteilte Annotationen ohne eigene Prüfung durch die Nutzenden (vorstrukturierend)	Alle Einsatzkontex-te	Sehr hohes Risiko. Normative Vorentscheidungen einer Person werden in fremden Verfahren wirksam;	Nicht empfohlen. Geteilte Annotationen nur, wenn jede nutzende Person sie einsehen, prüfen und anpassen kann und

Systemfunktion	Kontext	Risikobewertung	Bedingungen / Maßnahmen
		Meaningful Human Intervention strukturell gefährdet.	dies dokumentiert tut; Annotationsmonitoring (C.I.4.g.ee).

Die Stufenlogik verdeutlicht das zentrale Ergebnis der Evaluation: Das Risiko liegt nicht im MAKI-System als solchem, sondern in der Konfiguration aus Systemfunktion, Entscheidungskontext und institutioneller Einbettung. Dieselbe technische Infrastruktur kann in informierenden Funktionen und formalisierten Kontexten vertretbar und in empfehlenden Funktionen bei existenziellen Entscheidungsfolgen nur unter strengen Bedingungen akzeptabel sein. Die Entscheidung zwischen diesen Konfigurationen ist keine technische, sondern eine normative, und muss als solche transparent geführt werden.

Für MAKI 2.0 verschiebt sich das Bild, ohne dass sich die Grundstruktur der Bewertung ändert. Mit dem derzeitigen Verzicht auf Entscheidungsvorschläge entfällt die risikoreichste Funktion; bei der Textbausteinkomponente liegt die normative Entscheidung in den von Richterinnen und Richtern selbst angelegten Bausteinen. Zugleich verlagern sich die Risiken: Weil MAKI 2.0 auf großen Sprachmodellen beruht, treten an die Stelle der Ähnlichkeitsbeurteilung auf Grundlage vergangener Annotationen die für generative Systeme typischen Risiken der Konfabulation und der sprachlichen Glättung (Szenario 4). Sie betreffen vor allem Transformationsplatzhalter und Zusammenfassungen, deren Inhalt sich nicht unmittelbar auf eine Fundstelle in der Akte zurückführen lässt. Relationstabellen ordnen den Vortrag den Tatbestandsmerkmalen zu und sind nach der Unterscheidung des ersten Arbeitspakets eher als inhaltlich-normative Funktion einzuordnen. Die Bündelung gleichgelagerter Verfahren im Dashboard ist administrativ, kann aber standardisierend wirken, wenn Unterschiede zwischen den gebündelten Fällen aus dem Blick geraten. Die Governance-Frage geteilter Annotationen stellt sich beim „MAKItplace“ für geteilte Vorlagen und Textbausteine erneut. Schließlich bleibt die Grenze, die TabeA bewusst zieht, auch für künftige Anwendungen maßgeblich: Sobald das System selbst vorschlägt, welche Informationen als entscheidungsrelevant auszuwählen sind, kehrt das Risiko der Vorstrukturierung zurück.

Das MAKI-Beispiel illustriert auch eine risikokritische Dimension, die im fiktiven Strafzumessungsbeispiel nicht in dieser Form auftritt: die Governance-Frage der Annotationsausweitung. Der Übergang von individuellen zu geteilten Annotationen ist keine technische Systemfunktion, sondern eine institutionelle Entscheidung mit erheblichen normativen Konsequenzen. Sie erfordert eine explizite Regelung, die in der bestehenden Governance-Struktur des Projekts noch nicht vollständig abgebildet ist. Mit dem „MAKItplace“ stellt sie sich in MAKI 2.0 für geteilte Vorlagen und Textbausteine erneut.

Zugleich illustriert das Beispiel die Symmetriebedingung: Wenn ein gut gestaltetes, informierendes System nachweislich dazu beitragen kann, entscheidungsrelevante Informationen zuverlässiger und vollständiger bereitzustellen – etwa durch systematische Metadatenextraktion und Textvergleich in Fluggastrechteverfahren –, wäre der pauschale Verzicht auf seinen Einsatz seinerseits begründungsbedürftig. Die Entscheidung gegen KI-Nutzung ist ebenso bewertungsbedürftig wie die Entscheidung dafür.

5. Normative und institutionelle Konsequenzen

Die folgenden Ausführungen beziehen sich wiederholt auf die Pflichten der KI-Verordnung (Verordnung (EU) 2024/1689[53]) für Hochrisiko-KI-Systeme (insbesondere Art. 9, 10, 12, 14, 15, 26 und 72 KI-VO). Sie sind für Systeme einschlägig, die nach Anhang III Nr. 8 lit. a („Rechtspflege und demokratische Prozesse“) dazu bestimmt sind, Justizbehörden bei der Ermittlung und Auslegung von Sachverhalten und Rechtsvorschriften und bei der Anwendung des Rechts zu unterstützen. Die KI-Verordnung ist seit dem 1. August 2024 in Kraft; ihre Pflichten werden jedoch gestaffelt anwendbar. Durch die Verordnung (EU) 2026/1744[54] („Digital Omnibus on AI“) vom 8. Juli 2026 wurde der Anwendungsbeginn der Hochrisiko-Pflichten für Systeme nach Anhang III vom 2. August 2026 auf den 2. Dezember 2027 verschoben. Bereits anwendbar sind hingegen das Verbotsregime nach Art. 5 und die Vorgaben zur KI-Kompetenz nach Art. 4 (seit 2. Februar 2025, letztere in der durch die Verordnung (EU) 2026/1744 geänderten Fassung), die Pflichten für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck (seit 2. August 2025) sowie die Transparenzpflichten nach Art. 50 (seit 2. August 2026). Die folgende Analyse bezieht sich damit auf geltendes, in seinen Hochrisiko-Pflichten aber erst ab Dezember 2027 anwendbares Recht (de lege lata). Sie kann bereits jetzt als Orientierung für eine verantwortbare Systemgestaltung dienen.
Drei Vorschriften verdienen dabei besondere Aufmerksamkeit, weil sie unmittelbar an den Evaluationsrahmen anschließen. Erstens nimmt Art. 6 Abs. 3 KI-VO Systeme von der Hochrisiko-Einstufung aus, die das Ergebnis der Entscheidungsfindung nicht wesentlich beeinflussen, etwa weil sie nur eine eng gefasste Verfahrensaufgabe oder eine vorbereitende Aufgabe erfüllen. Nach Erwägungsgrund 61 gilt dasselbe für rein begleitende Verwaltungstätigkeiten. Die regulatorische Einordnung folgt damit derselben funktions- und kontextbezogenen Logik wie der hier entwickelte Rahmen (vgl. C.I.4.j). Zweitens verpflichtet Art. 27 KI-VO Betreiber, die Einrichtungen des öffentlichen Rechts sind, vor der Inbetriebnahme eines Hochrisiko-Systems nach Anhang III zu einer Grundrechte-Folgenabschätzung. Der Evaluationsrahmen aus Kapitel C.I.4 lässt sich als methodische Grundlage einer solchen Abschätzung verstehen. Drittens gewährt Art. 86 KI-VO betroffenen Personen ein Recht auf Erläuterung der Rolle des KI-Systems im Entscheidungsverfahren. Dieses Recht korrespondiert mit der Anforderung, dass die Begründung einer Entscheidung den tatsächlichen Einfluss des Systems erkennen lassen muss (vgl. C.I.4.f.dd, C.I.4.f.ee).

[53] Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (Verordnung über künstliche Intelligenz), ABl. L 2024/1689.
[54] Verordnung (EU) 2026/1744 des Europäischen Parlaments und des Rates vom 8. Juli 2026 zur Änderung der Verordnungen (EU) 2024/1689, (EU) 2018/1139 und (EU) 2023/1230 im Hinblick auf die Vereinfachung der Umsetzung harmonisierter Vorschriften für künstliche Intelligenz (Digital-Omnibus-Verordnung zur KI), ABl. L 2026/1744.

a. Verantwortungszuweisung im Kontext KI-gestützter richterlicher Entscheidungsfindung

aa. Das Responsibility-Gap-Problem

Die vorangegangene Analyse hat gezeigt, dass Risiken beim KI-Einsatz in der Justiz typischerweise aus dem Zusammenspiel technischer Systemgestaltung, individueller Faktoren und organisationaler Rahmenbedingungen entstehen. Diese Mehrebenenstruktur wirft die grundlegende Frage auf, wer die Verantwortung trägt, wenn eine KI-gestützte Entscheidung zu einem fehlerhaften Ergebnis führt.
In der philosophischen Debatte wurde dieses Problem als Responsibility Gap formuliert. Matthias[55] argumentierte als Erster, dass lernende, adaptive Systeme eine Situation erzeugen, in der weder Hersteller noch Betreiber das zukünftige Systemverhalten vollständig vorhersagen können und deshalb nach traditionellen Zurechnungskonzepten nicht verantwortlich gemacht werden können. Die Gesellschaft stehe vor der Wahl, entweder auf den Einsatz solcher Systeme zu verzichten oder eine Verantwortungslücke zu akzeptieren, die mit bestehenden Konzepten nicht geschlossen werden könne.
Santoni de Sio und Mecacci[56] haben diese Analyse weiterentwickelt und vier Dimensionen der Verantwortungslücke unterschieden: eine Culpability Gap, die entsteht, wenn die Ursache eines Schadens nicht eindeutig einem schuldhaft Handelnden zugeordnet werden kann, eine Moral Accountability Gap, wenn Systemverhalten weder erklärbar noch vorhersehbar ist, eine Public Accountability Gap, wenn öffentlich Verantwortliche ihr Handeln nicht mehr begründen und rechtfertigen können, sowie eine Active Responsibility Gap, wenn unklar ist, wer prospektiv für die Folgen eines Systemeinsatzes Sorge zu tragen hat.
Aus rechtlicher Sicht ist die Annahme, Verantwortung lasse sich nach herkömmlichen Regeln nicht zuweisen, für den Justizkontext nicht zutreffend. Im entscheidungsunterstützenden Einsatz bleibt die Entscheidungsverantwortung institutionell personengebunden (Art. 92, 97 GG) und an Begründungspflichten geknüpft; für Amtspflichtverletzungen haftet nach Art. 34 GG i. V. m. § 839 BGB der Staat. Eine rechtliche Verantwortungslücke besteht insoweit nicht. Die Problematik liegt an anderer Stelle: Eine algorithmische Vorstrukturierung von Entscheidungen (Szenarien 1–4) kann dazu führen, dass rechtliche Zuständigkeit und tatsächliche Kontrolle auseinanderfallen. Im Extremfall bleibt der Richter formal verantwortlich, während die Entscheidung faktisch durch Systemgestaltung, Trainingsdaten oder Interface-Design vorstrukturiert wurde. In den Kategorien von Santoni de Sio und Mecacci betrifft dies weniger die Culpability Gap als die Moral und die Public Accountability Gap: die Frage, ob der Richter eine Entscheidung noch begründen und rechtfertigen kann, deren Zustandekommen er nur eingeschränkt kontrolliert hat. Dieser Abschnitt behandelt Verantwortung daher nicht haftungsrechtlich, sondern fragt, wie Verantwortung zwischen den beteiligten Akteuren verteilt sein muss, damit rechtliche Zuständigkeit und tatsächliche Kontrolle nicht auseinanderfallen.

[55] *Matthias*, The responsibility gap: Ascribing responsibility for the actions of learning automata, Ethics and Information Technology 6 (2004), 175.
[56] *Santoni de Sio/Mecacci*, Four Responsibility Gaps with Artificial Intelligence: Why they Matter and How to Address them, Philosophy & Technology 34 (2021), 1057.

bb. Funktionale Verantwortungsverteilung

Eine pauschale Zuordnung der Verantwortung an einen einzelnen Akteur greift angesichts der Mehrebenenstruktur zu kurz. Die drei Evaluationsperspektiven (C.I.4.c, C.I.4.f, C.I.4.g) legen vielmehr eine **funktionale Verantwortungsverteilung** nahe, bei der jede Akteursgruppe für diejenigen Risiken verantwortlich ist, die in ihrem Einflussbereich entstehen und die sie durch angemessenes Handeln hätte verhindern oder begrenzen können.

Systemdesigner tragen Verantwortung für die technische Erfüllung der Aufsichtsbedingungen (C.I.4.d.bb). Dazu gehört die Gestaltung des Interfaces (epistemischer Zugang), die Vermeidung verhaltenslenkender Defaults (kausale Eingriffsmacht) und die transparente Kommunikation von Systemgrenzen. Ihre Verantwortung endet nicht mit der Auslieferung des Systems, sondern umfasst die kontinuierliche Überwachung, Rekalibrierung und die Bereitstellung von Gegenanzeigen (C.I.4.f.ff). Dieses Verständnis entspricht dem Anbieterbegriff der KI-Verordnung (Verordnung (EU) 2024/1689), deren Art. 14 Abs. 1 vorsieht, dass Hochrisiko-KI-Systeme so gestaltet werden müssen, dass sie von natürlichen Personen wirksam beaufsichtigt werden können.[57]

Die hier gewählte Terminologie weicht bewusst von den Kategorien der KI-Verordnung ab, die Verantwortung im Wesentlichen zwischen Anbieter und Betreiber verteilt – eine Zweiteilung, die etwa auch das BSI seinen Prüfkriterien zugrunde legt.[58] Der Evaluationsrahmen hebt demgegenüber die Anwenderebene als eigenständige dritte Kategorie hervor, weil Risiken, die aus der unmittelbaren Mensch-System-Interaktion entstehen (C.I.4.c–C.I.4.e), sich weder dem Anbieter- noch dem Betreiberbegriff vollständig zuordnen lassen. Systemdesigner und Organisation entsprechen damit im Kern dem Anbieter- und Betreiberbegriff der Verordnung; die Anwenderebene ist eine analytische Ausdifferenzierung des Betreiberbegriffs.

Organisationen (Justizverwaltungen, Gerichte) tragen Verantwortung für die institutionelle Ermöglichung wirksamer Kontrolle. Dazu gehört die Sicherung der Selbststeuerungsfähigkeit der Anwender (angemessene Arbeitsbedingungen, Zeitbudgets), die Gestaltung organisationaler Anreizstrukturen (keine impliziten Nutzungszwänge) und die Einrichtung von Monitoring- und Feedbackstrukturen (C.I.4.g.ee). Art. 26 Abs. 2 KI-VO konkretisiert dies für Betreiber: Sie müssen qualifiziertes Personal mit Kompetenz, Schulung und Befugnis für die Aufsicht einsetzen.

Anwender (Richterinnen und Richter) tragen Verantwortung für die Einzelfallentscheidung – aber nur unter der Bedingung, dass die Aufsichtsbedingungen (C.I.4.d.bb) tatsächlich erfüllt sind. Fehlt der epistemische Zugang, ist die Selbststeuerungsfähigkeit durch Arbeitsbelastung beeinträchtigt oder werden Abweichungen institutionell sanktioniert, kann die Verantwortung für eine fehlerhafte Übernahme der Systemempfehlung nicht allein dem Anwender zugeschrieben werden. Verantwortung setzt Kontrolle voraus.[59]

---

[57] *Fink*, Human Oversight under Article 14 of the EU AI Act, 2025, abrufbar unter https://doi.org/10.2139/ssrn.5147196 (Abruf am 25.10.2026, 14:02 Uhr).

[58] *Bundesamt für Sicherheit in der Informationstechnik (BSI)*, A5 – Horizontales Trustworthiness-Basismodul, Community-Draft, Version 0.9, Stand 30.06.2026, abrufbar unter https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KI/A5/A5_Horizontales-Trustworthiness-Basismodul.pdf (Abruf am 25.10.2026, 14:27 Uhr).

[59] Vgl. allgemein *Fischer/Ravizza*, Responsibility and Control – A Theory of Moral Responsibility, 1. Aufl. 1998; *Santoni de Sio/van den Hoven*, Meaningful Human Control over Autonomous Systems: A Philosophical Account, Frontiers in Robotics and AI 5 (2018), Art. 15.

cc. Bedingte Verantwortung und institutionelle Vorbedingungen

Die funktionale Verteilung impliziert ein Konzept **bedingter Verantwortung**. Die Verantwortung des Anwenders für die Letztentscheidung ist an die tatsächliche Erfüllung der Aufsichtsbedingungen geknüpft. Sind diese strukturell unterlaufen, beispielsweise durch asymmetrischen Begründungsaufwand (Szenario 2) oder fehlende epistemische Transparenz (Szenario 4), verschiebt sich die Verantwortung anteilig auf diejenigen Akteure, deren Handeln oder Unterlassen die Kontrolldefizite erzeugt haben.
Diese bedingte Verantwortung ist keine haftungsrechtliche Kategorie. Haftungsrechtlich bleibt die Zuweisung nach geltendem Recht möglich: Für Pflichtverletzungen im richterlichen Bereich haftet nach Art. 34 GG i. V. m. § 839 BGB der Staat, wobei das Spruchrichterprivileg (§ 839 Abs. 2 BGB) die Haftung für Urteile auf Straftaten beschränkt; für fehlerhafte Systeme kommt daneben die Produkthaftung des Herstellers in Betracht, die mit der Richtlinie (EU) 2024/2853[60] ausdrücklich auch Software erfasst. Das hier entwickelte Konzept setzt früher an: Es fragt nicht, wer für einen eingetretenen Schaden einzustehen hat, sondern wer die Voraussetzungen dafür schaffen muss, dass die rechtlich verantwortliche Person ihre Verantwortung tatsächlich wahrnehmen kann.
Eine tragfähige Verantwortungsverteilung muss daher die institutionellen Vorbedingungen wirksamer Kontrolle als eigenständige Verantwortungskategorie etablieren. Wer Systeme einsetzt, ohne die Aufsichtsbedingungen sicherzustellen, trägt eine institutionelle Mitverantwortung für Fehlentscheidungen, die aus diesem Kontrolldefizit resultieren. Die KI-Verordnung regelt diese Frage nicht: Art. 14 ist als Gestaltungspflicht der Anbieter konstruiert, die Zuweisung von Verantwortung zählt nicht zu seinen Zwecken.[61] Fink plädiert deshalb dafür, Verantwortung über den gesamten Lebenszyklus eines Systems zu verteilen; das stützt das hier vertretene Modell bedingter Verantwortung.

dd. Rückbindung an das Anwendungsbeispiel

Am Beispiel der KI-gestützten Strafzumessungsempfehlung (C.I.4.i) lässt sich die funktionale Verantwortungsverteilung konkretisieren. Wird das System als Default-Wert implementiert (dritte Stufe der Stufenlogik), trägt der Systemdesigner Verantwortung für die verhaltenslenkende Gestaltung, die Organisation für den institutionellen Kontext, der die Übernahme begünstigt, und der Richter für die Letztentscheidung – letzteres aber nur, soweit die Aufsichtsbedingungen tatsächlich erfüllt sind. Wird das System als Fallvergleichsanalyse implementiert (erste Stufe), verringert sich das Verantwortungsproblem erheblich, da der Automatisierungsgrad sinkt und die kausale Eingriffsmacht des Richters erhalten bleibt.
Die Stufenlogik aus C.I.4.i erweist sich damit auch als Heuristik der Verantwortungsallokation: Je höher der Automatisierungsgrad und je enger die faktische Entscheidungsfreiheit, desto stärker verlagert sich die Verantwortung von der Anwender- auf die Designer- und Organisationsebene.

[60] 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, ABl. L 2024/2853.
[61] *Fink*, Human Oversight under Article 14 of the EU AI Act, 2025, abrufbar unter https://doi.org/10.2139/ssrn.5147196 (Abruf am 25.10.2026, 14:02 Uhr).

b. Effizienz-Risiko-Abwägungen

aa. Struktur zulässiger Trade-offs

Der Evaluationsrahmen hat gezeigt, dass KI-Systeme in der richterlichen Entscheidungsfindung sowohl Effizienzgewinne als auch neue Risiken erzeugen können. Die Frage, ob und unter welchen Bedingungen ein Trade-off zwischen Effizienz und Risiko normativ akzeptabel ist, wurde bislang analytisch vorbereitet (Symmetriebedingung, Abschnitt C.I.2.d), aber nicht normativ beantwortet.
Grundsätzlich sind Trade-offs zwischen Effizienz und Risiko im Rechtssystem alltäglich. Jede Verfahrensbeschleunigung kann mit einem Verlust an Gründlichkeit verbunden sein, jede Standardisierung mit einem Verlust an Einzelfallsensibilität. Der KI-Einsatz erzeugt keine kategorial neue Abwägungssituation, verändert aber deren Struktur.
Einerseits verschiebt er die Sichtbarkeit von Risiken. Während die Effizienzgewinne algorithmischer Unterstützung häufig unmittelbar messbar sind (kürzere Bearbeitungszeiten, erhöhter Durchsatz), manifestieren sich die Risiken – insbesondere Automation Bias, schleichende Standardisierung und epistemische Fehlinterpretation – erst kumulativ und mit Verzögerung. Diese Asymmetrie begünstigt eine systematische Übergewichtung kurzfristiger Effizienzgewinne gegenüber langfristiger Risikosteigerung.[62]
Des Weiteren verändert der KI-Einsatz die Verteilung von Risiken und Gewinnen über die Akteursebenen. Effizienzgewinne kommen typischerweise der sekundären Ebene (Justizverwaltung) zugute, während Risiken – insbesondere epistemische Überbewertung und Fehlinterpretation – primär auf der primären Ebene (Verfahrensbeteiligte, Richter) wirksam werden. Dong et al.[63] haben für den Sozialleistungsbereich gezeigt, dass unmittelbar Betroffene Genauigkeitsverluste algorithmischer Entscheidungen deutlich weniger hinnehmen als weniger stark involvierte Akteure, und dass Letztere diese Differenz systematisch unterschätzen (vgl. Abschnitt C.I.2.c). Trade-offs, die Risiken auf die primäre Ebene verlagern, um Effizienzgewinne auf der sekundären Ebene zu realisieren, sind daher besonders begründungsbedürftig.
Letztlich entstehen mit zunehmender Skalierung qualitativ neue Risiken, die sich nicht als graduelle Erhöhung bestehender Risiken beschreiben lassen. Der Übergang vom idiosynkratischen zum systemischen Risiko (Szenario 3) bedeutet, dass ein lokaler Effizienzgewinn bei breiter Implementierung zu einem systemischen Kontrollverlust führen kann, der in keinem proportionalen Verhältnis zum ursprünglichen Gewinn steht. In der Netzwerkforschung wird dieses Phänomen als robust-yet-fragile-Muster beschrieben.[64]

bb. Unverhandelbare Schwellen

Nicht alle Risiken sind gegen Effizienzgewinne aufrechenbar. In Anlehnung an die rechtsphilosophische Unterscheidung zwischen abwägungsfähigen und abwägungsfesten Positionen lässt sich eine Klasse von Risiken identifizieren, bei denen ein Trade-off normativ unzulässig ist. Dies sind Risiken, die den Kern

[62] Vgl. allgemein *Kahneman*, Thinking, Fast and Slow, 1. Aufl. 2011.
[63] *Dong/Bonnefon/Rahwan*, Heterogeneous preferences and asymmetric insights for AI use among welfare claimants and non-claimants, Nature Communications 16 (2025), Art. 6973.
[64] *Haldane/May*, Systemic risk in banking ecosystems, Nature 469 (2011), 351.

richterlicher Unabhängigkeit betreffen. Wenn die Entscheidung im entscheidungs*unterstützenden* Kontext faktisch auf das System verlagert wird (Szenario 2: formale Kontrolle ohne faktische Eingriffsmacht), sind der Richtervorbehalt des Art. 92 GG und die richterliche Unabhängigkeit (Art. 97 GG) berührt; zudem kann die Garantie des gesetzlichen Richters (Art. 101 Abs. 1 S. 2 GG) betroffen sein. Kein Effizienzgewinn rechtfertigt eine institutionelle Konfiguration, in der menschliche Beteiligung zur legitimatorischen Geste reduziert wird.[65]
Wenn darüber hinaus ein System Empfehlungen generiert, die zu einer sachlich nicht gerechtfertigten Ungleichbehandlung anhand geschützter Merkmale führen, ist sein Einsatz unabhängig von sonstigen Effizienzgewinnen unzulässig. Dieser Grundsatz folgt aus den Diskriminierungsverboten (Art. 3 Abs. 3 GG, Art. 21 GRCh) und wird in der KI-Verordnung durch die Anforderungen an die Datengovernance (Art. 10 Abs. 2 lit. f, g) flankiert. Die Erfahrungen mit dem COMPAS-System in den USA haben gezeigt, dass algorithmische Risikobewertungen trotz formaler Neutralität systematische Ungleichheiten reproduzieren können.[66]
Ferner können diese Risiken die Transparenz und Begründbarkeit richterlicher Entscheidungen strukturell untergraben. Die fehlende Erklärbarkeit eines Systems schließt seinen Einsatz dabei nicht per se aus; dem Transparenzinteresse wird, wie bei menschlichen Entscheidungen, durch die nachträgliche Begründung der Entscheidung Rechnung getragen, die im entscheidungsunterstützenden Einsatz das Gericht selbst leistet (siehe zur sog. Black-Box-Problematik zudem die Abschnitte C.II.4.a und C.II.4.b.bb(1)).[67] Die Begründungspflicht (Art. 20 Abs. 3, Art. 103 Abs. 1 GG; Art. 6 EMRK) verlangt jedoch, dass das Gericht seine Entscheidung aus den rechtlichen und tatsächlichen Umständen des Einzelfalls heraus selbst tragen kann. Eine Konfiguration, in der Systemempfehlungen übernommen werden, ohne dass das Gericht sie in dieser Weise eigenständig begründen könnte, ist daher nicht durch Effizienzgewinne kompensierbar. Fallbezogene Erklärbarkeit des Systems (C.I.4.f.ee) ist hierfür keine rechtliche Voraussetzung, erleichtert dem Gericht aber die eigenständige Prüfung und Begründung erheblich.
Diese Einschätzung wird durch die bereichsspezifische Perspektive der CEPEJ gestützt: Deren Guidelines on the Use of Generative AI for Courts formulieren mit dem Grundsatz, dass generative Outputs niemals bindend sein dürfen und der Zugang zu einer menschlichen richterlichen Entscheidung stets gewährleistet bleiben muss, im Kern dieselbe unverhandelbare Schwelle, die hier aus Art. 92, 97 GG und der KI-Verordnung abgeleitet wird[68] – unabhängig von deren jeweiligem Anwendungsbereich. (Siehe umfassend zu den *rechtlichen* roten Linien des Einsatzes informationstechnischer Systeme bei der richterlichen Entscheidungsfindung den Abschnitt C.II.)

[65] *Green*, The flaws of policies requiring human oversight of government algorithms, Computer Law & Security Review 45 (2022), Art. 105681.
[66] *Carnat*, Addressing the risks of generative AI for the judiciary: The accountability framework(s) under the EU AI Act, Computer Law & Security Review 55 (2024), Art. 106067.
[67] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 149f.
[68] *European Commission for the Efficiency of Justice (CEPEJ)*, Guidelines on the use of generative AI for courts, CEPEJ(2025)18Final, 2025, abrufbar unter https://rm.coe.int/cepej-2025-18final-en-draft-guidelines-on-the-use-of-generative-ai-for/48802a4ad1 (Abruf am 25.10.2026, 14:29 Uhr).

cc. Entscheidungskompetenz über Trade-offs

Wo Trade-offs normativ zulässig sind – also außerhalb der unverhandelbaren Schwellen –, stellt sich die Frage, wer über ihre Akzeptabilität entscheidet.
Die Festlegung der Rahmenbedingungen – welche Systemfunktionen in welchen Verfahrensarten grundsätzlich zulässig sind, welche Mindestanforderungen gelten und welche Ausschlussbereiche bestehen – ist eine Aufgabe des Gesetzgebers und der Justizverwaltung. Sie erfordert eine demokratisch legitimierte Entscheidung, die die in C.I.5.b.bb identifizierten Schwellen verbindlich festschreibt.
Die Umsetzung im Einzelfall – ob ein System in einer konkreten Entscheidungssituation genutzt wird, welches Gewicht seiner Empfehlung beigemessen wird und wann eine Abweichung geboten ist – liegt beim richterlichen Anwender. Diese Kompetenz ist Ausdruck richterlicher Unabhängigkeit und kann weder durch institutionelle Effizienzvorgaben noch durch technische Defaults übersteuert werden.
Die technische Ausgestaltung der Abwägungsoptionen liegt bei den Systemdesignern. Sie müssen sicherstellen, dass die Stufenlogik (C.I.4.i) technisch abgebildet ist: dass also ein System, das als informierendes Werkzeug konzipiert ist, nicht durch Gestaltungsentscheidungen faktisch zum empfehlenden oder vorstrukturierenden System wird. Verantwortbare Systemgestaltung schließt die Bereitschaft ein, die eigenen Produkte dort zu begrenzen, wo ihr Einsatz mit den identifizierten Risikoprofilen nicht vereinbar ist (C.I.4.f.ff).

c. Moderatoren und Schutzmaßnahmen

aa. Systematik der Maßnahmen

Die Risikoszenarien (C.I.4.e) und das Anwendungsbeispiel (C.I.4.i) haben exemplarisch gezeigt, welche konkreten Maßnahmen den identifizierten Risikopfaden entgegenwirken können.
Maßnahmen zur Risikoreduktion lassen sich in drei Klassen einteilen[69]: **Insulation** (Vermeidung der Risikosituation), **Mitigation** (Abschwächung der Risikowirkung) und **Removal Strategies** (Beseitigung der Risikoursache). Diese Einteilung entspricht im Kern den Optionen der Risikobehandlung nach ISO 31000.[70] Die Wahl der geeigneten Klasse hängt dabei vom Risikoprofil ab. Bei hohem Risiko und unverhandelbaren Schwellen (C.I.5.b.bb) kommt nur Insulation oder Removal in Betracht, bei moderatem Risiko können Mitigationsmaßnahmen ausreichen.
Die Forschung zu Debiasing-Strategien in professionellen Entscheidungskontexten verlagert den Schwerpunkt zunehmend vom Individuum auf die Gestaltung des Entscheidungsumfelds. Eine systematische Übersicht, überwiegend aus dem medizinischen Bereich, ergab, dass 80 % der untersuchten technologischen Strategien wirksam waren, wobei informationsdesignbasierte Ansätze am häufigsten untersucht wurden.[71] Für den Justizkontext schlussfolgern Próchnicki et al.[72], dass

[69] *Stafford*, Biases in decision making, Tribunals, Winter 2017, 19.
[70] *International Organization for Standardization (ISO)*, ISO 31000:2018 – Risk management – Guidelines, Genf 2018, Abschn. 6.5.2.
[71] *Dharanikota/Howie/Hope/Wigmore/Skipworth/Yule*, Debiasing Judgements Using a Distributed Cognition Approach: A Scoping Review of Technological Strategies, Human Factors 67 (2025), 525.
[72] *Próchnicki/Janik/Bystranowski*, Debiasing Numerical Verdicts and Judicial Discretion: Reflections on Mitigating the Anchoring Effect in Judicial Decision-Making, in: Bystranowski/Janik/Próchnicki (Hrsg.),

Debiasing-Maßnahmen im Kontext richterlicher Ermessensausübung besonders sorgfältig kalibriert werden müssen, um nicht ihrerseits legitime Entscheidungsspielräume einzuschränken. (Siehe auch zu Maßnahmen im Umgang mit der Bias-Thematik insbesondere die Abschnitte C.II.2.a.bb(3) sowie C.II.2.c.bb(3))

bb. Technische Maßnahmen (Designerebene)

Technische Maßnahmen adressieren primär die Risikopfade der epistemischen Überbewertung (Szenario 1) und der Fehlinterpretation (Szenario 4). Sie operieren auf der Ebene der Interface-Gestaltung und der Systemarchitektur.

(1) Informationsdesign und Unsicherheitskommunikation

Ein Ansatzpunkt gegen epistemische Überbewertung ist die kontextualisierte Darstellung von Systemoutputs. Dies umfasst statt einer numerischen Punktschätzung die Angabe von Bandbreiten und der Größe der Referenzklasse. Krauss et al.[73] fanden beispielsweise in einer Simulation, dass die kommunizierte Risikostufe nur bei kategorialer, nicht aber bei probabilistischer Darstellung auf die Entscheidung durchschlug; Green & Chen[74] zeigten in einem Experiment mit Laien für algorithmische Risikoprognosen, dass Darstellungsbedingungen das Verhalten der Entscheidenden veränderten.
Die Wirksamkeit sollte aber nicht überschätzt werden. Green und Chen fanden zugleich, dass die Probanden weder die Genauigkeit ihrer eigenen noch die der algorithmischen Prognose zuverlässig einschätzen konnten. Und Gravett[75] argumentiert für den Justizkontext, dass bloße Warnhinweise – im Fall State v. Loomis ein vorgeschriebener Hinweis auf die Grenzen des COMPAS-Scores – wirkungslos bleiben, weil sie den Anker unverändert lassen und der Automation Bias sie überspielt. Die Darstellungsform kann also beeinflussen, welches Gewicht eine Empfehlung erhält, sie kann aber die Überschätzungstendenz nicht aufheben. Ein Warnhinweis neben einem unveränderten Score ist deshalb keine Maßnahme.
Die KI-Verordnung setzt an der schwächeren Stelle an: Art. 14 Abs. 4 lit. b verlangt lediglich, dass Anwender in die Lage versetzt werden, die eigene Tendenz zur automatischen Übernahme zu erkennen. Eine Anforderung an die Darstellung des Outputs selbst enthält die Norm nicht.

(2) Technische Maßnahmen gegen Konfabulation

Für generativ arbeitende Systemkomponenten treten spezifische technische Gegenmaßnahmen neben die beschriebenen Informationsdesign-Maßnahmen. Retrieval-Augmented-Generation-Ansätze, bei denen ein System Antworten auf

Judicial Decision-Making – Integrating Empirical and Theoretical Perspectives, 1. Aufl., Cham 2022, S. 79ff.

[73] *Krauss/Cook/Klapatch*, Risk assessment communication difficulties: An empirical examination of the effects of categorical versus probabilistic risk communication in sexually violent predator decisions, Behavioral Sciences & the Law 36 (2018), 532.

[74] *Green/Chen*, Disparate Interactions: An Algorithm-in-the-Loop Analysis of Fairness in Risk Assessments, Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* ’19), 2019, 90.

[75] *Gravett*, Judicial Decision-Making in the Age of Artificial Intelligence, in: Sousa Antunes/Freitas/Oliveira/Martins Pereira/Vaz de Sequeira/Barreto Xavier (Hrsg.), Multidisciplinary Perspectives on Artificial Intelligence and the Law, 1. Aufl., Cham 2024, S. 281ff.

Grundlage einer geprüften Wissensbasis – etwa justizinterner Datenbanken – generiert, statt sie allein aus gelernten Mustern zu erzeugen, senken das Konfabulationsrisiko spürbar, beseitigen es jedoch nicht: Auch RAG-gestützte Modelle bewegen sich in der oben genannten Bandbreite unbelegter Aussagen. Ergänzend kommen die Zerlegung von Ausgaben in einzeln gegen die Wissensbasis prüfbare Teilaussagen sowie Konsistenzprüfungen mehrfach erzeugter Antworten auf dieselbe Anfrage in Betracht: Widersprechen sich unabhängig generierte Ausgaben zur selben Anfrage, deutet dies auf eine Konfabulation hin. Keine dieser Maßnahmen ersetzt die in C.I.4.f.aa geforderte Quellenkennzeichnung und die menschliche Prüfung im Einzelfall.

(3) Funktionale Beschränkung

Wie das Anwendungsbeispiel (C.I.4.i) gezeigt hat, ist die wirksamste einzelne Maßnahme die Reduktion des Automatisierungsgrades: von empfehlend auf informierend. Durch die Beschränkung auf eine Fallvergleichsanalyse anstelle einer konkreten Empfehlung sinkt das Risikoprofil erheblich, ohne dass die informativen Teilfunktionen verloren gehen. Diese Maßnahme ist eine Form der Insulation: Sie vermeidet die Risikosituation, statt sie zu bekämpfen.

(4) Mindest-Benchmarks als Ausschlusskriterien

Wird dennoch an einer Empfehlungsfunktion festgehalten, müssen Mindestanforderungen definiert werden, bei deren Nichterfüllung der Einsatz unzulässig ist. Dazu gehören: dokumentierte Modellgüte auf repräsentativen Testdaten, Fairness-Audits bezüglich geschützter Merkmale sowie regelmäßige Rekalibrierung bei veränderter Rechtslage. Für die Modellgüte entstehen inzwischen domänenspezifische Prüfverfahren für das deutsche Recht, etwa der an der TU München entwickelte Benchmark BenGER, der Sprachmodelle anhand subsumtionsbasierter Falllösungen bewertet.[76] Solche Benchmarks messen allerdings die juristische Leistungsfähigkeit an Ausbildungs- und Prüfungsaufgaben; ob ein System in einer konkreten Verfahrensart, etwa im Fluggastrechte- oder Asylverfahren, zuverlässig arbeitet, muss zusätzlich anhand von Daten aus diesem Einsatzkontext geprüft werden. Art. 9 KI-VO verlangt ein Risikomanagementsystem, das solche Anforderungen operationalisieren kann. Für die konkrete Ausgestaltung dieser Mindestanforderungen bietet sich ein Rückgriff auf bereits bestehende oder in Entstehung befindliche Kriterienkataloge an. Das BSI operationalisiert in seinem A5-Basismodul entsprechende Qualitätsmanagementprozesse für Transparenz, Robustheit, Performance, Bias und Erklärbarkeit[77], jeweils mit getrennten Anforderungen an Anbieter und Betreiber. Ein Rückgriff auf einen solchen Katalog vermeidet, dass die hier vorgeschlagenen Mindest-Benchmarks freihändig statt entlang eines in Entstehung befindlichen behördlichen Prüfstandards entwickelt werden.

---

[76] *Nagl/Mayrhofer/Heidebach/Koçak/Zettelmeier/Breu/Greiner/Milijas/Grabmair*, BenGER: Benchmarking LLM Systems on Subsumption-Based Legal Reasoning in German Law, 2026, abrufbar unter https://arxiv.org/abs/2605.28183 (Abruf am 25.10.2026, 14:05 Uhr).

[77] *Bundesamt für Sicherheit in der Informationstechnik (BSI)*, A5 – Horizontales Trustworthiness-Basismodul, Community-Draft, Version 0.9, Stand 30.06.2026, GCS.3.1–3.6, abrufbar unter https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KI/A5/A5_Horizontales-Trustworthiness-Basismodul.pdf (Abruf am 25.10.2026, 14:27 Uhr).

(5) Symmetrischer Begründungsaufwand

Eine technische Maßnahme gegen den Risikopfad der eingeschränkten Intervention (Szenario 2) besteht darin, den Begründungsaufwand für Übernahme und Abweichung anzugleichen. Wenn Abweichungen von Systemempfehlungen begründet werden müssen, muss auch die Übernahme einer Empfehlung kurz begründet werden. Damit wird die asymmetrische Begründungslast neutralisiert, die den Kern des zweiten Risikopfades bildet. Dieser Vorschlag setzt nicht voraus, dass das System schlechter urteilt als das Gericht allein. Auch bei erwiesener Leistungsfähigkeit bleibt die Umkehrung der Begründungslast rechtsstaatlich problematisch, weil sie die formale Entscheidungsverantwortung von der faktischen Entscheidungsmacht entkoppelt.

cc. Organisationale Maßnahmen

Organisationale Maßnahmen adressieren primär die Risikopfade der eingeschränkten Intervention (Szenario 2) und der institutionellen Standardisierung (Szenario 3).

(1) Cognitive Forcing Mechanisms

Die organisational bedeutsamste Maßnahme ist die verpflichtende Eigeneinschätzung vor Einsicht in die Systemempfehlung. Der Richter formuliert zunächst eine eigene vorläufige Bewertung; erst danach wird die Systemeinordnung sichtbar. Damit wird der Ankereffekt strukturell unterbrochen. Buçinca et al.[78] haben in einem Online-Experiment mit Laien geprüft, ob prozedurale Unterbrechungen des Entscheidungsflusses die Übernahme fehlerhafter Systemempfehlungen verringern. In einer der Bedingungen mussten die Probanden zunächst selbst entscheiden, bevor die Systemempfehlung sichtbar wurde. Cognitive Forcing reduzierte die Übernahmerate deutlich stärker als das bloße Anzeigen von Erklärungen. Jedoch wurden gerade die wirksamsten Varianten von den Nutzern am schlechtesten bewertet, und der Effekt fiel bei Personen mit höherer Bereitschaft zu kognitiver Anstrengung stärker aus. Für den Justizkontext stellen Guthrie et al.[79] fest, dass auch erfahrene Richter kognitiven Verzerrungen unterliegen und dass institutionelle Gegenmaßnahmen individueller Reflexion überlegen sind.

(2) Monitoring mit Frühwarnschwellen

Die systematische Erhebung von Übernahmequoten, Override-Raten und Entscheidungsvarianz ermöglicht die Früherkennung schleichender Standardisierung (Szenario 3). Definierte Schwellenwerte, etwa eine Übernahmequote über 90 % in einem Spruchkörper, lösen eine institutionelle Prüfung aus. Art. 72 (Anbieter) und Art. 26 Abs. 5 KI-VO (Betreiber) verlangen eine Beobachtung im Betrieb, die um solche Kennzahlen ergänzt werden kann. Solche Kennzahlen bewegen sich allerdings in einem verfassungsrechtlich sensiblen Bereich. Die Dienstaufsicht über Richterinnen und Richter findet einfachgesetzlich nach § 26 Abs. 1 DRiG nur statt, soweit sie die richterliche Unabhängigkeit nicht beeinträchtigt; der Kernbereich der Rechtsprechung,

[78] *Buçinca/Malaya/Gajos*, To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making, Proceedings of the ACM on Human-Computer Interaction 5 (2021), CSCW1, Art. 188.
[79] *Guthrie/Rachlinski/Wistrich*, Judging by Heuristic: Cognitive Illusions in Judicial Decision Making, Judicature 86 (2002), 44.

zu dem auch die Entscheidung über Nutzung und Gewichtung einer Systemempfehlung gehört, ist ihr entzogen. Unzulässig sind auch Maßnahmen, die mittelbar Druck auf den Inhalt richterlicher Entscheidungen ausüben können. Eine Übernahmequote, die auf der Ebene eines Spruchkörpers erhoben und mit einer institutionellen Prüfung verknüpft wird, kann diese Grenze berühren, zumal in Verfahren, in denen der Spruchkörper faktisch aus einer einzelnen Person besteht, wie beim Einzelrichter im Asylverfahren oder beim Amtsgericht. Monitoring-Kennzahlen dienen daher ausschließlich der Bewertung des Systems und seiner institutionellen Einbettung, nicht der Bewertung richterlicher Arbeit. Sie sollten so aggregiert werden, dass Rückschlüsse auf einzelne Richterinnen und Richter ausgeschlossen sind, und ihre Auswertung sollte unter Beteiligung der richterlichen Vertretungen geregelt werden.

(3) Institutionell geschützte Nichtnutzung

Die Entscheidung gegen den Systemeinsatz darf keinem institutionellen Kostendruck unterliegen. Art. 14 Abs. 4 lit. d KI-VO kodifiziert dieses Recht: Anwender müssen in der Lage sein, das System in einer konkreten Situation nicht zu nutzen, seinen Output zu ignorieren oder eine davon abweichende Entscheidung zu treffen. Die institutionelle Verankerung erfordert, dass individuelle Nutzungsmuster nicht statistisch überwacht und Nichtnutzungen nicht als Qualitätsmangel interpretiert werden.

(4) Schulung mit Schwerpunkt auf Systemgrenzen

Kompetenzaufbau muss über die Vermittlung technischer Funktionsweisen hinausgehen und systematisch behandeln: die Differenz zwischen Gruppen- und Individualprognose, bekannte Bias-Risiken, Automation Bias und den epistemischen Status statistischer Empfehlungen. Stafford[80] betont, dass Anti-Bias-Strategien nur dann nachhaltig wirken, wenn sie institutionell verankert und nicht auf individuelle Reflexion beschränkt sind. Die KI-Verordnung unterstützt dies durch Art. 4, der Anbieter und Betreiber zu Maßnahmen zur Förderung der KI-Kompetenz ihres Personals verpflichtet (AI Literacy).

dd. Verfahrensbezogene Maßnahmen

Verfahrensbezogene Maßnahmen operieren auf der Ebene des Entscheidungsprozesses selbst und adressieren übergreifend alle vier Risikopfade.

(1) Externe Audits

Regelmäßige unabhängige Überprüfungen der Systemleistung, der Nutzungsmuster und der Entscheidungsqualität ergänzen das interne Monitoring. Diese Empfehlung geht bewusst über die KI-Verordnung hinaus. Zwar verlangt Art. 9 Abs. 2 KI-VO ein Risikomanagementsystem als kontinuierlichen, iterativen Prozess über den gesamten Lebenszyklus eines Systems. Für Hochrisiko-Systeme nach Anhang III Nr. 8 genügt jedoch eine Konformitätsbewertung auf Grundlage interner Kontrolle durch den Anbieter selbst (Art. 43 Abs. 2 i. V. m. Anhang VI); eine unabhängige Stelle wird nicht eingebunden. Die Prüfung der Systeme, die in die richterliche Entscheidungsfindung

[80] *Stafford*, Biases in decision making, in: Tribunals, Winter 2017, 19.

eingreifen, läge nach der Verordnung damit in erster Linie bei denjenigen, die sie entwickeln. Gerade im Justizkontext, in dem Anbieter häufig privatwirtschaftlich organisiert sind (vgl. C.I.2.c) und die Wirkungen eines Systems erst im Zusammenspiel mit Nutzungspraxis und Organisation sichtbar werden (C.I.4.h), erscheint eine externe Überprüfung als notwendige Ergänzung. Als Grundlage können bestehende oder entstehende Prüfstandards dienen, etwa ISO/IEC 42001:2023 oder der Entwurf des A5-Basismoduls des BSI (vgl. C.I.6.b). Gegenstand solcher Audits sind das System und seine institutionelle Einbettung, nicht die Entscheidungen einzelner Richterinnen und Richter; ihre Ausgestaltung muss die richterliche Unabhängigkeit wahren.

(2) Fehleranalyse als systemische Untersuchung

Wenn Fehlentscheidungen auftreten, dürfen sie nicht ausschließlich dem einzelnen Richter zugeschrieben werden. Root-Cause-Analysen müssen die Mehrebenenstruktur berücksichtigen: War das Interface irreführend gestaltet? Waren die Arbeitsbedingungen angemessen? War die Schulung ausreichend? Dieser Ansatz entspricht der in der Sicherheitsforschung etablierten Unterscheidung zwischen individuellen Fehlern und systemischen Ursachen.[81]

(3) Transparente Dokumentation der Systemnutzung

Für eine spätere Überprüfbarkeit ist zu dokumentieren, ob und in welchem Umfang ein KI-System in die Entscheidungsfindung einbezogen wurde, welche Empfehlung das System generiert hat und ob der Entscheidungsträger davon abgewichen ist. Diese Dokumentation dient nicht der Überwachung individueller Richter, sondern der institutionellen Lernfähigkeit und der Wahrung des Rechts auf eine nachvollziehbare Entscheidungsbegründung. Art. 12 KI-VO verlangt eine automatische Protokollierung, die diese Anforderung technisch unterstützen kann (siehe zu den Protokollierungspflichten nach Art. 12 KI-VO zudem Abschnitt C.II.4.b.cc(2)).

ee. Zuordnung zu Risikopfaden

Die folgende Übersicht ordnet die wichtigsten Maßnahmen den vier Risikopfaden aus C.I.4.d.dd zu:

Risikopfad	Technische Maßnahmen	Organisationale Maßnahmen	Verfahrensbezogene Maßnahmen
1. Epistemische Überbewertung	Informationsdesign, Unsicherheitskommunikation; Funktionale Beschränkung (informierend statt empfehlend)	Cognitive Forcing Mechanisms (Eigeneinschätzung vor Systemeinsicht); Schulung zu Systemgrenzen	—
2. Eingeschränkte Intervention	Symmetrischer Begründungsaufwand; Frühe Interventionspunkte;	Institutionell geschützte Nichtnutzung; Keine Nutzungsquoten	—

[81] Vgl. allgemein *Reason*, Human Error, 1. Aufl. 1990.

	Kein vorausgefüllter Default	als Leistungsindikator	
3. Institutionelle Standardisierung	—	Monitoring mit Frühwarnschwellen; Rollenklarheit und Ausschlussbereiche; Keine impliziten Nutzungserwartungen	Externe Audits; Systematische Auswertung von Übernahmequoten und Entscheidungsvarianz
4. Fehlinterpretation	Fallbezogene Erklärbarkeit; Gegenanzeigen; Dokumentierte Modellgrenzen	Schulung zur Differenz Gruppen-/Individualprognose; Interdisziplinäre Trainings	Transparente Dokumentation der Systemnutzung; Protokollierung gem. Art. 12 KI-VO

Diese Zuordnung zeigt, dass kein einzelner Risikopfad durch eine einzelne Maßnahme auf einer einzelnen Ebene wirksam unterbrochen werden kann. Risikoreduktion erfordert ein koordiniertes Zusammenwirken technischer, organisationaler und verfahrensbezogener Maßnahmen. Dies entspricht dem zentralen Befund des Evaluationsrahmens: Risiken entstehen aus dem Zusammenspiel aller drei Ebenen und können nur durch ein Zusammenwirken aller drei Ebenen wirksam begrenzt werden.

ff. Verhältnismäßigkeit der Maßnahmen

Der Aufwand der Schutzmaßnahmen muss zum Risikoprofil der Anwendung passen. Ein informierendes System mit moderatem Risikoprofil erfordert epistemische Transparenz, Schulung und Monitoring, aber keine Cognitive Forcing Mechanisms oder externe Audits. Ein empfehlendes System mit hohem Risikoprofil erfordert das volle Maßnahmenspektrum einschließlich Mindest-Benchmarks und Frühwarnschwellen. Eine vorstrukturierende Konfiguration ist – im Bereich der Strafzumessung – durch Maßnahmen nicht auf ein akzeptables Risiko zu reduzieren; hier ist Nichtnutzung die risikoadäquate Maßnahme.
Dieser Verhältnismäßigkeitsmaßstab entspricht dem Grundsatz der KI-Verordnung, dass Aufsichtsmaßnahmen den Risiken, dem Grad der Autonomie und dem Nutzungskontext des Systems angemessen sein müssen (Art. 14 Abs. 3).

6. Schlussbetrachtung

a. Ertrag der Analyse

Der vorliegende Bericht hat einen Evaluations- und Reflexionsrahmen für den Einsatz von KI-Systemen in der richterlichen Entscheidungsfindung entwickelt. Sein zentrales Ergebnis lässt sich in einem Satz zusammenfassen: Das Risiko liegt nicht im System als solchem, sondern in der Konfiguration aus Systemfunktion, Entscheidungssituation und institutioneller Einbettung.
Diese Einsicht hat konkrete analytische Konsequenzen. Der Evaluationsrahmen operationalisiert sie über drei Risikodimensionen (Normativität, Automatisierungsgrad, Komplexität), vier Aufsichtsbedingungen (kausale Eingriffsmacht, epistemischer Zugang, Selbststeuerungsfähigkeit, passende Absichten), vier typische Risikopfade (epistemische Überbewertung, eingeschränkte Intervention, institutionelle Standardisierung, Fehlinterpretation) und drei Evaluationsperspektiven (Anwender, Systemdesigner, Organisation). Die durchlaufenden Anwendungsbeispiele – die fiktive Strafzumessungsempfehlung und das reale System MAKI – haben gezeigt, dass diese Instrumente eine differenzierte Bewertung ermöglichen, die weder in pauschaler Ablehnung noch in unkritischer Übernahme mündet, sondern in einer Stufenlogik, die den Einsatz an kontextspezifische Bedingungen knüpft.
Die normativen Konsequenzen in Kapitel C.I.5 haben drei Ergebnisse hervorgebracht: ein Konzept funktionaler und bedingter Verantwortungsverteilung, das Verantwortung an die tatsächliche Erfüllung der Aufsichtsbedingungen koppelt, die Identifikation unverhandelbarer Schwellen, bei denen kein Effizienzgewinn einen Trade-off rechtfertigt, und einen systematisierten Maßnahmenkatalog, der technische, organisationale und verfahrensbezogene Interventionen den vier Risikopfaden zuordnet.

b. Grenzen des Rahmens

Es soll betont werden, dass der Evaluationsrahmen für einen spezifischen Anwendungskontext entwickelt wurde: KI-Systeme, die richterliche Entscheidungen informieren, strukturieren oder empfehlen, innerhalb des institutionellen Rahmens der deutschen und europäischen Justiz.
Als zentrale Limitation ist zu nennen, dass der Rahmen auf qualitativen Risikoszenarien, nicht auf empirischen Wirkungsstudien beruht. Die vier Risikopfade sind theoretisch plausibel und durch Forschung aus angrenzenden Domänen gestützt, aber nicht durch kontrollierte Studien im Justizkontext validiert. Ob und in welchem Ausmaß Automation Bias, institutionelle Standardisierung oder epistemische Fehlinterpretation tatsächlich auftreten, ist eine empirische Frage, die der Rahmen stellen, aber nicht beantworten kann.
Diese Lücke ließe sich gerade bei bereits erprobten Systemen schließen. Für MAKI und TabeA kämen insbesondere in Betracht: eine Befragung der erprobenden Richterinnen und Richter zu Nutzung, Vertrauen und Abweichungsverhalten; eine Auswertung der Systemprotokolle nach Art. 12 KI-VO zu Übernahmequoten und Entscheidungsvarianz, aggregiert im Sinne von C.I.5.c.cc; eine Prüfung der Quelltreue von TabeA-Zusammenfassungen gegen die zugrunde liegenden Anhörungsprotokolle; sowie Vignettenstudien mit Richterinnen und Richtern zur Ankerwirkung von Systemvorschlägen. Erst solche Daten würden den in C.I.4.b geforderten Vergleich mit dem Status quo ermöglichen.
Ferner fokussiert sich der Rahmen auf die primäre Akteursebene und die Mensch-System-Interaktion. Infrastrukturelle Risiken (Cyberangriffe, Plattformabhängigkeiten,

Datenlecks), die in Kapitel C.I.2 als eigenständige Risikokategorie eingeführt wurden, werden im Evaluationsrahmen nur selektiv berücksichtigt. Für eine vollständige Risikobewertung müssten diese Ebenen ergänzend analysiert werden.
Diese Selbstbeschränkung betrifft insbesondere die technischen Phasen des KI-Lebenszyklus vor und nach der Entscheidungsinteraktion: Datenakquisition und -validierung, sicherheitsbezogene Verifikation, Deployment-Governance sowie Retirement bzw. Außerbetriebnahme eines Systems. Für diese Dimensionen existieren bereits spezialisierte, prüfbare Kontrollkataloge – etwa der Entwurf des Horizontalen Trustworthiness-Basismoduls der KI-Prüfarchitektur A5 des Bundesamts für Sicherheit in der Informationstechnik[82], das Anforderungen entlang des gesamten Lebenszyklus getrennt für Anbieter und Betreiber formuliert. International ergänzt der zertifizierbare Standard ISO/IEC 42001:2023 für KI-Managementsysteme dieses Bild.[83] Der vorliegende Evaluationsrahmen versteht sich als komplementär zu einem solchen technischen Prüfrahmen: Er verortet Risiken im Entscheidungskontext, ersetzt aber keine technische Auditierung der zugrundeliegenden Systeme.
Letztlich ist der Rahmen statisch konzipiert, obwohl er Iterativität fordert. Er beschreibt, welche Fragen zu stellen sind und welche Maßnahmen in Betracht kommen, aber er kann die Dynamik technologischer Entwicklung, sich verändernder Nutzungspraktiken und institutioneller Lernprozesse nur postulieren, nicht abbilden. Diese Grenze ist nicht nur ein methodisches Defizit. Sie verweist auf ein strukturelles Problem, das über den Rahmen selbst hinausgeht.

c. Das Regulierungsdilemma: Laufen, um zu stehen

„Now, here, you see, it takes all the running you can do, to keep in the same place.“ – The Red Queen, in Lewis Carroll, Through the Looking-Glass (1871)

Die zuletzt genannte Grenze des Rahmens ist zugleich seine politisch folgenreichste. Risikodimensionen, Aufsichtsbedingungen und Risikopfade wurden bewusst technologieneutral entwickelt und gelten unabhängig davon, ob ein System klassifiziert, prognostiziert oder generiert (vgl. C.I.4.c.ff sowie die Szenarien 1 und 4 in C.I.4.e). Verschoben hat sich in letzter Zeit vor allem die konkrete Praxis: Große Sprachmodelle und multimodale Systeme haben sich rasant weiterentwickelt, und die Standards zur Messung und Kommunikation ihrer spezifischen Risiken – etwa von Konfabulation – sind erkennbar jünger und weniger gefestigt als die etablierten Verfahren für klassifikatorisch-prognostische Systeme. Die strukturell tiefere Verschiebung betrifft jedoch nicht den Wechsel von Klassifikation zu Generierung, sondern den Wechsel von einem einzelnen, prüfbaren Output zu selbständig verketteten Handlungsschritten – zunehmend autonome Agenten.
Die Europäische Union hat mit der KI-Verordnung (Verordnung (EU) 2024/1689) das weltweit erste umfassende Regulierungswerk für künstliche Intelligenz vorgelegt. Ihre Architektur – risikobasierte Klassifikation, Anforderungen an Transparenz, Datengovernance und menschliche Aufsicht – ist konzeptionell kohärent und normativ ambitioniert. Art. 14 kodifiziert wesentliche Elemente des hier entwickelten

[82] *Bundesamt für Sicherheit in der Informationstechnik (BSI)*, A5 – Horizontales Trustworthiness-Basismodul, Community-Draft, Version 0.9, Stand 30.06.2026, abrufbar unter https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KI/A5/A5_Horizontales-Trustworthiness-Basismodul.pdf (Abruf am 25.10.2026, 14:27 Uhr).
[83] *International Organization for Standardization/International Electrotechnical Commission (ISO/IEC)*, ISO/IEC 42001:2023 – Information technology – Artificial intelligence – Management system, Genf 2023.

Evaluationsrahmens: die Forderung nach wirksamer menschlicher Aufsicht, die Berücksichtigung von Automation Bias, das Recht auf Nichtnutzung.
Zugleich offenbart die Analyse ein strukturelles Dilemma. Die Regulierung muss so schnell laufen, wie sie kann, um nicht zurückzufallen. Die Halbwertszeit regulatorischer Annahmen über Systemarchitekturen, Interaktionsmuster und Risikoprofile schrumpft mit jeder Modellgeneration. Die KI-Verordnung wurde auf Grundlage eines technologischen Stands entworfen, der bei ihrem Inkrafttreten bereits teilweise überholt war. Die Einfügung von Bestimmungen zu Allzweck-KI-Modellen (General Purpose AI, Art. 51 ff.) im Trilog-Verfahren illustriert dieses Muster: Ein regulatorisches Rahmenwerk, das für aufgabenspezifische Hochrisikosysteme konzipiert war, musste im letzten Moment um eine Kategorie erweitert werden, die bei der ursprünglichen Konzeption noch nicht im Fokus stand.
Ein noch aktuelleres Beispiel liefert die Verordnung (EU) 2026/1744 vom 8. Juli 2026 („Digital Omnibus on AI“), die den Anwendungsbeginn der Hochrisiko-Pflichten aus Anhang III – und damit auch der hier einschlägigen Kategorie Rechtspflege – von August 2026 auf Dezember 2027 verschoben hat. Die Union verschiebt damit den Anwendungsbeginn der eigenen Pflichten, noch bevor diese überhaupt gegriffen haben, für genau jene Systeme, die im Zentrum dieses Berichts stehen – ein Beleg dafür, dass das Wettlaufmuster nicht erst bei der technologischen Anpassung ansetzt, sondern bereits die Umsetzung eines als fertig geglaubten Rahmenwerks betrifft.
Dieses Hinterherhinken ist kein Implementierungsdefizit, sondern ein Strukturmerkmal technologischer Regulierung unter Bedingungen schneller, z. T. exponentieller, Innovationszyklen.
Der Rahmen dieses Berichts steht unter Vorbehalten, die außerhalb seiner Reichweite liegen. Der erste betrifft die Geschwindigkeit technologischer Veränderung. Zwischen dem Kommissionsentwurf der KI-Verordnung im April 2021 und ihrem Inkrafttreten im August 2024 haben ChatGPT, GPT-4, Gemini, Claude und zahlreiche offene Modelle die Leistungsgrenzen generativer Systeme mehrfach verschoben. Die Risikoszenarien der Verordnung beruhen auf Fähigkeitsannahmen, die heute konservativ wirken. Ein Rahmen, der für klassifikatorische Risikobewertungssysteme entwickelt wurde, muss für generative Systeme erweitert werden, und diese Erweiterung wird von der nächsten Modellgeneration eingeholt.
Hinzu kommt eine geopolitische Asymmetrie. Die leistungsfähigsten Modelle entstehen überwiegend in den USA und in China, unter grundlegend anderen regulatorischen Bedingungen. In den USA dominieren sektorale Regulierung und freiwillige Selbstverpflichtung, in China verbindet sich staatliche Kontrolle mit aggressiver Innovationsförderung, einschließlich des Einsatzes von KI im Justizsystem unter Bedingungen, die den hier entwickelten Anforderungen nicht entsprechen. Die KI-Verordnung gilt zwar für jedes in der EU eingesetzte System unabhängig vom Herstellungsort. Über Modellarchitektur, Trainingsdaten und Sicherheitsmaßnahmen wird aber außerhalb ihres Zugriffs entschieden. Die EU reguliert damit den Einsatz von Systemen, deren Entstehung sie nur mittelbar beeinflusst.
Auf multilateraler Ebene versucht die UNESCO mit ihren Draft Guidelines for the Use of AI Systems in Courts and Tribunals[84] einen globalen Mindeststandard zu setzen – ein Versuch, der die beschriebene Asymmetrie normativ adressiert, ohne sie auflösen zu können.
Daraus erwächst ein Regulierungsrisiko eigener Art. Wo technologische Überlegenheit als geopolitischer Faktor gilt, gerät regulatorische Ambition unter

[84] *UNESCO*, Guidelines for the Use of AI Systems in Courts and Tribunals, Paris 2025, abrufbar unter https://doi.org/10.58338/LIEY8089 (Abruf am 04.11.2026, 20:45 Uhr).

Rechtfertigungsdruck. Die Balance zwischen Schutz und Wettbewerbsfähigkeit ist kein stabiles Gleichgewicht, sondern ein politisch permanent zu verteidigendes Ergebnis. Der Grundsatz unverhandelbarer Schwellen aus C.I.5.b.bb steht unter dem ständigen Druck einer Argumentation, die Effizienz gegen Grundrechtsschutz aufrechnet.
Am weitesten reicht jedoch eine Verschiebung, die der Rahmen dieses Berichts noch nicht abbildet. Die bisherige Analyse unterstellt Systeme, die einen Output erzeugen, den ein Mensch bewertet. Agentische Systeme, die Arbeitsschritte selbständig verketten, also recherchieren, auswerten, Entwürfe erzeugen und Ergebnisse weiterverarbeiten, verändern diese Grundannahme. Je mehr Schritte zwischen menschlicher Freigabe und Ergebnis liegen, desto später wird ein Fehler sichtbar und desto mehr Folgeschritte hat er bereits geprägt. Die Aufsichtsbedingungen aus C.I.4.d.bb verlieren dabei nicht graduell, sondern strukturell an Wirkung. Epistemischer Zugang setzt einen prüfbaren Einzeloutput voraus, kausale Eingriffsmacht einen Zeitpunkt, zu dem ein Eingriff noch etwas ändert. Auch Art. 14 Abs. 4 lit. e KI-VO denkt den Betrieb als anhaltbaren Vorgang, nicht als Kette bereits vollzogener Handlungen. Die Stufenlogik aus C.I.4.i wäre insoweit um eine vierte Stufe zu ergänzen, deren Risikoprofil sich nicht durch bessere Kontrolle des Einzelfalls senken lässt, sondern nur durch Begrenzung des Handlungsspielraums.

d. Konsequenzen für die Justiz

Für den spezifischen Kontext der richterlichen Entscheidungsfindung folgen aus diesem Regulierungsdilemma entscheidende Konsequenzen.
Die erste Konsequenz ist methodischer Natur. Der hier entwickelte Evaluationsrahmen kann nicht als einmalig anzuwendendes Instrument verstanden werden. Er muss als iterativer Prozess institutionalisiert werden, der mit jeder relevanten technologischen Veränderung erneut durchlaufen wird. Die Kontextanalyse (C.I.4.c), die Prüfung der Aufsichtsbedingungen (C.I.4.d) und die Risikoszenarien (C.I.4.e) müssen an neue Systemarchitekturen angepasst werden. Die Monitoring- und Feedbackstrukturen (C.I.4.g.ee) sind nicht nur Schutzmaßnahmen, sondern der institutionelle Mechanismus, über den diese Anpassung erfolgt. Sie sind der Punkt, an dem der Evaluationsrahmen selbstreflexiv wird: Er bewertet nicht nur Systeme, sondern überwacht die Angemessenheit seiner eigenen Bewertungskriterien.
Die zweite Konsequenz betrifft die institutionelle Infrastruktur. Die Justiz benötigt eigene Kompetenz zur Bewertung von KI-Systemen. Die in C.I.4.g.dd und C.I.5.c.cc geforderten Schulungs- und Qualifizierungsmaßnahmen reichen nicht aus, wenn das Wissen über Systemfähigkeiten, -grenzen und -risiken ausschließlich von den Systemanbietern selbst stammt. Unabhängige, justizinterne oder justiznahe Evaluationskapazitäten – etwa in Form spezialisierter Prüfstellen, interdisziplinärer Beratungsgremien oder kooperativer Forschungsstrukturen zwischen Rechtswissenschaft, Informatik und Kognitionspsychologie – sind eine Voraussetzung dafür, dass die Justiz nicht zum passiven Abnehmer technologischer Entwicklungen wird, sondern deren Bedingungen aktiv mitgestaltet.
In einer Welt, in der die leistungsfähigsten KI-Systeme unter Bedingungen entwickelt werden, die europäischen Grundrechtsstandards nicht entsprechen, wird die Frage, welche Systeme in der europäischen Justiz eingesetzt werden dürfen, zu einer Entscheidung über das Verhältnis von technologischer Effizienz und rechtsstaatlicher Integrität. Der vorliegende Bericht hat argumentiert, dass diese Entscheidung nicht technisch determiniert ist, sondern normativ geführt werden muss. Die unverhandelbaren Schwellen – richterliche Unabhängigkeit, Diskriminierungsfreiheit, Begründbarkeit – sind keine Hindernisse, die überwunden werden müssen, sondern

die Bedingungen, unter denen der Einsatz von KI in der Justiz überhaupt legitimierbar ist.

e. Ausblick

Die Rote Königin in Lewis Carrolls Through the Looking-Glass beschreibt eine Welt, in der man so schnell laufen muss, wie man kann, um an derselben Stelle zu bleiben. Für die Regulierung von KI im Justizkontext ist dieses Bild zutreffend, aber nicht vollständig. Denn die Frage ist nicht nur, ob die Regulierung mit der technologischen Entwicklung Schritt halten kann. Die Frage ist, ob sie das überhaupt muss.
Ein Regulierungsansatz, der darauf abzielt, jede neue Systemarchitektur vorab vollständig zu erfassen, wird scheitern. Die Halbwertszeit technologiespezifischer Normen ist zu kurz. Erfolgversprechender ist ein Rahmen, der auf stabilen Prinzipien beruht und deren Operationalisierung flexibel an neue Kontexte anpasst. Die KI-Verordnung verfolgt diesen Ansatz im Grundsatz. Ihre Anforderungen an Transparenz, menschliche Aufsicht und Risikomanagement sind technologieneutral formuliert. Der vorliegende Evaluationsrahmen übersetzt diese Prinzipien in kontextspezifische Reflexionsfragen, die unabhängig von der konkreten Systemarchitektur anwendbar sind. Die Frage, ob ein System die vier Aufsichtsbedingungen erfüllt, ist für ein Risikobewertungstool ebenso relevant wie für einen generativen Schreibassistenten – auch wenn die Antworten unterschiedlich ausfallen.
Die eigentliche Aufgabe liegt demnach nicht in der Vorhersage technologischer Entwicklungen, sondern in der Sicherung institutioneller Lernfähigkeit. Die Justiz muss in die Lage versetzt werden, neue Systeme anhand bewährter Prinzipien zu bewerten, anstatt für jedes neue System neue Prinzipien entwickeln zu müssen.
Die geopolitische Dimension verstärkt diese Dringlichkeit, ohne den Grundsatz zu verändern. Dass die leistungsfähigsten Systeme außerhalb europäischer Regulierungskontexte entstehen, ist kein Argument dafür, europäische Standards abzusenken. Es ist ein Argument dafür, sie als Einsatzbedingungen verbindlich durchzusetzen und die institutionelle Kapazität aufzubauen, die dafür erforderlich ist. Eine Justiz, die alternativ Systeme einsetzt, deren Wirkungsweise sie nicht versteht, deren Risiken sie nicht bewerten kann und deren Folgen sie nicht verantwortet, wäre nicht ein Verlust an Wettbewerbsfähigkeit. Sie wäre ein Verlust an Rechtsstaatlichkeit.
Der Wettlauf der Roten Königin ist nicht zu gewinnen, wenn das Ziel darin besteht, jede technologische Entwicklung regulatorisch einzuholen. Er ist zu gewinnen, wenn die Prinzipien, die den Lauf orientieren, stabil genug sind, um auch unter veränderten Bedingungen zu tragen. Richterliche Unabhängigkeit, Diskriminierungsfreiheit und die Begründbarkeit von Entscheidungen sind solche Prinzipien. Sie sind nicht verhandelbar, nicht weil sie einer vergangenen Welt angehören, sondern weil sie die Bedingung dafür sind, dass technologischer Wandel in der Justiz als Fortschritt und nicht als Schwächung erfahren wird.

II. Rechtliche Handlungsoptionen und Handlungsempfehlungen

1. Einleitung

Aus den geltenden verfassungs-, menschen-, europa- und zivilprozessrechtlichen Rahmenbedingungen ergeben sich zahlreiche rote Linien für den Einsatz informationstechnischer Systeme bei der richterlichen Entscheidungsfindung. Dies ist ausführlich in dem ersten Arbeitspaket (Milestone 1) des Forschungsprojekts „Künstliche Intelligenz und richterliche Entscheidungsfindung“ herausgearbeitet worden.[85] Auf Grundlage dieser Ergebnisse werden nun Handlungsoptionen untersucht und diskutiert, die einen diese Linien wahrenden Einsatz informationstechnischer Systeme bei der richterlichen Entscheidungsfindung ermöglichen. Aus diesen Handlungsoptionen wird jeweils eine Handlungsempfehlung (i. S. e. „Gold-Standards“) herausgearbeitet, die die Verwirklichung der rechtlichen Anforderungen bei dem Einsatz informationstechnischer Systeme in der richterlichen Entscheidungsfindung am besten sicherstellt.
Die identifizierten roten Linien lassen sich thematisch clustern, d. h. sie können **übergeordneten Themenfeldern**, wie beispielsweise dem Erfordernis menschlicher Kontrolle und Aufsicht (siehe dazu Abschnitt C.II.2), dem Erfordernis von Transparenz und Information (siehe dazu Abschnitt C.II.4) bzw. dem Erfordernis einer Rechtsgrundlage für den Einsatz informationstechnischer Systeme (siehe dazu Abschnitt C.II.6) zugeordnet werden. Der vorliegende Berichtsteil nimmt eine entsprechende Clusterung in den Gliederungsabschnitten C.II.2 bis C.II.9 vor und führt die Erkenntnisse des ersten Arbeitspakets (Milestone 1) somit thematisch zusammen.
Darüber hinaus lassen sich die identifizierten roten Linien **kategorisieren**. Es sind einerseits solche roten Linien zu identifizieren, die übergreifend für *jeden* (d. h. sowohl den entscheidungs*ersetzenden als auch den* entscheidungs*unterstützenden*) Einsatz eines informationstechnischen Systems gelten. Andererseits existieren rote Linien, die jeweils nur für die eine oder die andere Kategorie, d. h. nur für den entscheidungs*ersetzenden* Einsatz bzw. nur für den entscheidungs*unterstützenden* Einsatz heranzuziehen sind. Dieses Dokument greift diese Kategorisierung auf und unterscheidet in den jeweiligen thematischen Clustern anhand dieser Kategorisierung.
Im Vorgriff zu den nachfolgenden inhaltlichen Ausführungen des vorliegenden Arbeitspakets soll Folgendes festgehalten werden: Mit Blick auf eine schnell spürbare Entlastung der Gerichte sollte beachtet werden, dass insbesondere die Leistungsfähigkeit informationstechnischer Systeme im Bereich der Verarbeitung und Strukturierung großer Datenmengen erhebliche Vorteile bei der schnellen Erledigung von Verfahren, die umfangreiche Datenmengen beinhalten, verspricht (z. B. in Massenverfahren wie Dieselverfahren und in Verfahren bei Verletzungen von Regelungen der DSGVO, in denen Schriftsätze mehrere hundert Seiten umfassen).[86] Es sollten daher Fallgestaltungen für die Implementierung informationstechnischer Systeme bei der richterlichen Entscheidungsfindung priorisiert anvisiert werden, bei

[85] Siehe hierzu ausführlich *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025.
[86] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 115, vgl. auch die *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, Münchener Thesen zum Zivilprozess der Zukunft, abrufbar unter https://www.olg-duesseldorf.nrw.de/infos/Zivilprozess-der-Zukunft/Muenchener-Thesen-zum-Zivilprozess-der-Zukunft.pdf (Abruf am 26.02.2026, 18:27 Uhr), S. 9, These B.3.

denen eine Entscheidung auf eine Vielzahl von Fällen angewendet werden kann.[87] Derartige Verfahren sollten prioritär – mit Blick auf die Überlastung der Gerichte – identifiziert werden und sodann primär für diese Verfahren informationstechnische Systeme zur Entlastung der menschlichen Richterinnen und Richter entwickelt werden.[88] Dies verspricht eine Kapazitätsoptimierung in doppelter Hinsicht: Die eingesetzten informationstechnischen Systeme würden in diesen – für menschliche Richterinnen und Richter ressourcenintensiven – Verfahrenskategorien ungleich schneller Ergebnisse erzeugen und die Verfahren zu einem Abschluss bringen können. In der Folge können die ansonsten für diese Verfahren gebundenen Richterinnen und Richter die weiteren, ihnen obliegenden Verfahren zu einem zügigeren Abschluss bringen.[89] Eine schrittweise Einführung – auch unter Beachtung der Thematik der Bestandsinfrastruktur – kann hierbei eine bereits frühzeitig ansetzende, sukzessiv ansteigende Entlastungswirkung bedingen, gegenüber einer Einführung aller erdenklichen Funktionen am Ende des vollständigen Entwicklungsabschlusses.[90]
Die Funktionsfähigkeit des Rechtsstaats ist von kaum zu überschätzender Bedeutung. Informationstechnische Systeme bei der richterlichen Entscheidungsfindung können ihren Beitrag hierzu leisten.

[87] *Horn*, ZZP 136 (2023), 459, 481; in Bezug auf entscheidungsunterstützende Systeme vgl. auch Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 102f.
[88] Siehe zur Kapazitäts(ü)be(r)lastung der Gerichte auch in anderen Staaten *Turbék*, Justiz und KI, Friedrich-Naumann-Stiftung für die Freiheit (Hrsg.), 2026, abrufbar unter https://shop.freiheit.org/#!/Publikation/2184 (Abruf am 09.10.2026, 15:40 Uhr), S. 13; vgl. auch *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 27f.
[89] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 117f.
[90] *Turbék*, Justiz und KI, Friedrich-Naumann-Stiftung für die Freiheit (Hrsg.), 2026, abrufbar unter https://shop.freiheit.org/#!/Publikation/2184 (Abruf am 09.10.2026, 15:40 Uhr), S. 30f. empfiehlt ebenfalls eine schrittweise Einführung, indes anhand des Risikoumfangs der Anwendungen; *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 23f.

2. Menschliche Aufsicht und Kontrolle des Systems

Die Notwendigkeit der Aufsicht und Kontrolle des informationstechnischen Systems bei der richterlichen Entscheidungsfindung durch menschliche Richterinnen und Richter ist eine der zentralen roten Linien, deren Einhaltung es zu gewährleisten gilt.

a. Übergreifende rote Linien für jeden Einsatz sowie Handlungsleitlinien

aa. Übergreifende rote Linien

Aus mehreren Verbürgungen lässt sich ein generelles Erfordernis von Aufsicht und Kontrolle des informationstechnischen Systems ableiten, das jeweils unterschiedliche Ausprägungen annimmt.

(1) Umfassende Einbindung menschlicher Richterinnen und Richter

Erstens ergibt sich das Erfordernis von Aufsicht und Kontrolle dahingehend, dass **menschliche Richter und Richterinnen** in die Beschaffung, den Betrieb und die Überwachung des informationstechnischen Systems **umfassend einzubinden** sind.[91] Soweit entscheidungsunterstützende Funktionen den Kernbereich der Rechtsprechung berühren (insbesondere inhaltsbezogene Funktionen) sowie im Falle des entscheidungsersetzenden Einsatzes bezieht sich die Notwendigkeit einer Einbindung auch auf die Entwicklung des informationstechnischen Systems.[92]
Dies ergibt sich insbesondere aus dem Demokratieprinzip[93], der Gewaltenteilung gemäß Art. 20 Abs. 2 Satz 2, Abs. 3 GG[94], den Art. 92[95], 97[96], 101 Abs. 1 Satz 2 GG[97], aus Art. 6 Abs. 1 EMRK[98] und aus Art. 47 Abs. 1, Abs. 2 GRCh[99]. Auch lässt sich dies aus Art. 26 Abs. 2, Abs. 5 UAbs. 1 KI-VO bzw. aus Art. 14 KI-VO für Hochrisiko-KI-Systeme ableiten.[100]
Eine Beauftragung der Entwicklung des informationstechnischen Systems durch Private oder die Verwaltung ist dabei grundsätzlich möglich, soweit menschliche Richter und Richterinnen hinreichend Einfluss auf die Entwicklung, das Training (und insbesondere die Auswahl, Gestaltung und Kontrolle der Trainingsdaten sowie der Entscheidungsbäume), die Konfiguration und den Betrieb des informationstechnischen Systems nehmen und dessen Funktionsfähigkeit

[91] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 126f., 131ff., 195ff., 245ff., 259ff.
[92] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 126f., 131ff., 168f., 195ff., 245f., 259ff.
[93] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 126f.
[94] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 132f.
[95] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 195ff.
[96] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 195ff.
[97] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 219ff.
[98] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 241ff.
[99] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 267.
[100] Ausführlich zu Anforderungen und Grenzen der Europäischen KI-Verordnung siehe *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 269ff.

überwachen sowie kontrollieren können.[101] Auch ist im Falle einer Überlassung des Betriebs des informationstechnischen Systems an Dritte eine inhaltliche Kontrolle richterlicher Dokumente durch diese zwingend auszuschließen.[102]

(2) Ausschluss von Bias

Zweitens erfordert Aufsicht und Kontrolle gleichzeitig eine inhaltliche Überwachung der Arbeit des Systems, um Fehlausgaben, die auf ***Bias***[103] beruhen, auszuschließen. Ein entscheidungsersetzender sowie -unterstützender Einsatz scheidet immer dann aus, wenn nicht sichergestellt werden kann, dass kein systemseitiger ***Bias*** besteht oder ein solcher sich nicht auf das Ergebnis auswirkt, sowie eine hinreichende Kontrolle durch einen menschlichen Richter oder eine menschliche Richterin gewährleistet ist.[104]
Dies ergibt sich insbesondere aus der Menschenwürdegarantie gemäß Art. 1 Abs. 1 GG[105] und Art. 1 GRCh[106], aus Art. 6 Abs. 1 EMRK[107] und Art. 47 Abs. 1, Abs. 2 GRCh[108], aus dem Gleichbehandlungsgrundsatz gemäß Art. 3 GG[109], Art. 14, 6 Abs. 1 EMRK[110] sowie Art. 20–26 GRCh[111] und aus dem Anspruch auf ein faires Verfahren gemäß Art. 2 Abs. 1, 3 Abs. 1, 20 Abs. 3 GG[112]. Zudem wird die Unparteilichkeit der entscheidenden Instanz von Art. 92, 97, 101 Abs. 1 Satz 2 GG[113] gewährleistet. Für Hochrisiko-KI-Systeme ergibt sich dies beispielsweise mit Blick auf Eingabedaten aus Art. 26 Abs. 4 KI-VO.[114]
Insofern ist auf eine diskriminierungssensible Auswahl der Trainingsdaten und eine menschliche Kontrolle dieser Trainingsdaten sowie der Arbeitsergebnisse des informationstechnischen Systems zu achten.[115]
Steht ein *Bias* zu befürchten, ist dies aus zivilprozessualer Sicht strukturell der Besorgnis der Befangenheit eines menschlichen Richters oder einer menschlichen Richterin vergleichbar.[116] Insoweit sind bereits systemseitige Vorkehrungen zu treffen,

[101] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 90f., 126f., 131ff., 195ff. (s. a. insb. 201f.), 221f., 241f., 245ff., 259ff., 267, 359.
[102] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 131f., 197f., 221f., 245f., 267.
[103] Schreibweise nach *Duden*, abrufbar unter https://www.duden.de/rechtschreibung/Bias (Abruf am 30.07.2026, 14:08 Uhr).
[104] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 89ff., 108ff., 169ff., 199ff., 219ff., 241ff., 260ff., 267, 356ff.
[105] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 89ff.
[106] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 267.
[107] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 241ff.
[108] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 267.
[109] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 108ff.
[110] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 246ff., 252ff., 255f.
[111] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 267.
[112] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 356ff.
[113] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 169ff., 199ff., 219ff.
[114] Ausführlich zu Anforderungen und Grenzen der Europäischen KI-Verordnung siehe *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 269ff.
[115] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 96ff., 108ff., 131ff., 195ff., 202, 216ff., 241ff., 254ff., 265ff., 356ff.
[116] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 265ff., 355ff.

um die Freiheit des Systems von *Bias* sicherzustellen und den ordnungsgemäßen sowie fairen Ablauf des Verfahrens sicherzustellen.[117]

bb. Handlungsoptionen und -empfehlungen

(1) Richterliche Aufsicht und Kontrolle bei Beschaffung, Betrieb, Überwachung und Entwicklung

(i) Handlungsoptionen

Möglichkeiten, richterliche Aufsicht und Kontrolle bei der Beschaffung, dem Betrieb, der Überwachung und in den relevanten Konstellationen auch bei der Entwicklung sicherzustellen, bestehen vielfach.
Eine **erste, durchaus sinnvolle Handlungsoption** besteht darin, die Beschaffung, Entwicklung, den Betrieb und die laufende Überwachung des informationstechnischen Systems in ein dauerhaftes **richterliches Governance-Modell** einzubetten.[118] Die dauerhafte Ausgestaltung entspricht hierbei auch einer zugrundeliegenden Wertung der KI-VO. So weist Art. 26 KI-VO untypischerweise Betreiberpflichten aufgrund des **dynamischen Charakters von KI-Systemen** auf, im Sinne einer wirksamen Überwachung zur Risikominimierung für Grundrechte und Sicherheit.[119] Die konkrete Ausgestaltung dieses dauerhaften richterlichen Governance-Modells ist eine primär organisatorische Frage, auch mit Blick auf die Entscheidung, wo der Betrieb des Systems angesiedelt sein soll. Der Betrieb kann prinzipiell bei der Judikative selbst, bei der Exekutive bzw. bei (privaten) Dritten liegen.[120] Maßgeblich i. S. e. Governance-Modells und mit Blick auf die vorstehend genannte rote Linie ist, dass die organisatorische Anbindung dergestalt erfolgt, dass richterliche Aufsicht und Kontrolle durch menschliche Richterinnen und Richter nicht nur punktuell, sondern vielmehr institutionell abgesichert über den gesamten Lebenszyklus des Systems erfolgt.[121] In der Literatur werden hierzu diverse **Handlungsvarianten** diskutiert:
Erstens kann die organisatorische Einbettung etwa durch die Einrichtung eines *interdisziplinär besetzten, richterlich(dominiert)en Steuerungs- oder Fachgremiums* geschehen, etwa in Form von richterlichen, sog. „Automatisierungskammern“.[122] Diese könnten u. a. für bestimmte Fall-Kategorien Rechtsfragen abstrakt (d. h. nicht bezogen auf einen konkreten Einzelfall) wie herkömmliche Spruchkörper entscheiden und hierüber das informationstechnische System (je nach Systemart im Rahmen eines Entscheidungsbaums bzw. mittels KI-Trainings) inhaltlich gestalten sowie dessen Betrieb laufend überwachen (dies kann z. B. auch die Aufsicht umfassen, dass das System auf den/die menschliche/n Benutzer/in reagiert; dies ist insbesondere mit Blick auf Art. 26 Abs. 2 KI-VO und Art. 14 Abs. 3 lit. b KI-VO bei Hochrisiko-KI-Systemen

[117] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 106ff., 167ff., 198, 219ff., 255f., 261, 265ff., 355f.
[118] Vgl. etwa *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 368ff.
[119] BeckOK KI-Recht/*Denga*, 7. Ed. 1.8.2026, KI-VO Art. 26 Rn. 1ff.; Martini/Wendehorst/*I. Eisenberger/Th. Binder*, 2. Aufl. 2026, KI-VO Art. 26 Rn. 1f.
[120] *Horn*, ZZP 136 (2023), 459, 478.
[121] Vgl. *Horn*, ZZP 136 (2023), 459, 478f.
[122] *Horn*, ZZP 136 (2023), 459, 478f.

relevant)[123]. Die vom System nicht lösbaren Fälle, die zu Aussteuerungen[124] an menschliche Richterinnen und Richter führen, könnten von diesen Kammern bearbeitet werden und die aus diesen Fällen gezogenen Erkenntnisse könnten durch diese Kammern genutzt werden, um das System fortentwickeln.[125] Hierbei wird vorgeschlagen, dass das Gremium nicht nur mit rechtswissenschaftlich ausgebildeten Richterinnen und Richtern besetzt ist, sondern gerade mit Blick auf Fragen der Programmierung oder das Design der Anwendung auch mit Expertinnen und Experten anderer Fachrichtungen, insbesondere der Informatik und (Arbeits- und Organisations-)Psychologie, die als unabhängige Richterinnen und Richter berufen werden könnten. Hierüber würde dem Gremium die für die Systementwicklung notwendige interdisziplinäre Fachkompetenz verliehen werden.[126] Gleichwohl sind auch andere Ausgestaltungen eines solchen Gremiums möglich, in denen die nicht-juristischen Disziplinen in bloß beratender bzw. ausführender Funktion, d. h. ohne Stimmrecht, eingebunden sind.[127] Neben einer Programmierung im herkömmlichen Sinne kann hierbei insbesondere die Aufsetzung eines sog. „No Code-Systems“ anvisiert werden, die ohne tiefgreifende Programmierkenntnisse ausgebaut werden können und mit Blick auf ihre Sicherheit auch unabhängig zertifiziert werden könnten.[128] Eine einzelne Kammer wird angesichts des präsumtiven Aufwands freilich überfordert sein.[129] Eine Handlungsoption könnte daher der Aufbau mehrerer Automatisierungskammern bei höherinstanzlichen Gerichten sein. Dies erscheint aufgrund der dort vorherrschenden hohen rechtlichen Expertise sinnvoll.

Als weitere Handlungsvariante wäre der Aufbau einer – auch bundesländerübergreifenden – *Automatisierungsinstitution* denkbar, etwa eines spezifischen Automatisierungsgerichts.[130] Dem Grunde nach kann hierbei auch als weitere Handlungsvariante der kumulative Aufbau der genannten Strukturen erfasst werden, indem beispielsweise die Automatisierungsinstitution die grundlegenden, eher technischen, Aspekte wahrnimmt und die an den Einzelgerichten bestehenden Automatisierungskammern die Aussteuerungen bearbeiten und das Feedback an die Automatisierungsinstitution leiten.

Neben dem vorstehend skizzierten Modell im Eigenausbau durch die Justiz kommt als weitere Variante die *institutionelle Arbeit im Verbund mit Externen* (Exekutive oder [private] Dritte), die das System bereitstellen und in Kooperation mit Richterinnen und

---

[123] Zudem können hiernach auch ggf. besondere informierende bzw. beratende Mechanismen des KI-Systems selbst notwendig sein. BeckOK KI-Recht/*Denga*, 7. Ed. 1.8.2026, KI-VO Art. 26 Rn. 32, 34; Martini/Wendehorst/*I. Eisenberger/Th. Binder*, 2. Aufl. 2026, KI-VO Art. 26 Rn. 26; BeckOK KI-Recht/*Buchner*, 7. Ed. 1.8.2026, KI-VO Art. 14 Rn. 42f.

[124] Siehe zur Aussteuerung *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 90, 145, 230, 292, 335, 344, 347; Mit dem Begriff der „Aussteuerung“ wird beschrieben, dass ein informationstechnisches System in der Konstellation des entscheidungsersetzenden Einsatzes der Bearbeitung eines Falles bzw. der Entscheidung über einen Fall enthoben wird und dieser Fall einem menschlichen Richter bzw. einer menschlichen Richterin zur Bearbeitung bzw. Entscheidung zugewiesen wird.

[125] *Horn*, ZZP 136 (2023), 459, 478f., 486.

[126] Vgl. *Horn*, ZZP 136 (2023), 459, 478f.; vgl. auch *Turbék*, Justiz und KI, Friedrich-Naumann-Stiftung für die Freiheit (Hrsg.), 2026, abrufbar unter https://shop.freiheit.org/#!/Publikation/2184 (Abruf am 09.10.2026, 15:40 Uhr), S. 31.

[127] Vgl. *Wolff*, Algorithmen als Richter, 1. Aufl. 2022, S. 182.

[128] *Horn*, ZZP 136 (2023), 459, 479; vgl. auch *Köhler*, Richterliche Verantwortung und Künstliche Intelligenz, 1. Aufl. 2026, S. 254.

[129] *Horn*, ZZP 136 (2023), 459, 478f.

[130] In eine ähnliche Richtung gehend, indem er eine „hochrangige Wertekommission“ beschreibt *Rollberg*, Algorithmen in der Justiz, 1. Aufl. 2020, S. 131; *Wolff*, Algorithmen als Richter, 1. Aufl. 2022, S. 182 fordert ein staatliches, unabhängiges Algorithmen-Gremium, bei dem jedoch die nicht-juristische Expertise in beratender oder ausführender Funktion ohne eigenes Stimmrecht ausgestattet ist.

Richtern (weiter)entwickeln, fachliche Anforderungen definieren, Testkriterien festlegen, Freigaben für neue Funktionen erteilen und das System fortlaufend begleiten, in Betracht.[131] Derartige Partnerschaften sowie Gremien aus Experten verschiedener Disziplinen sind bereits gelebte Praxis, etwa mit Blick auf die dem vorliegenden Forschungsprojekt entspringende MAKI-KI, aber auch mit Blick auf den Oberlandesgerichtsassistenten „OLGA“, der die Richterinnen und Richter der Diesel-Senate des OLG Stuttgart unterstützt. Entsprechende Verbünde versprechen größere finanzielle und personelle Flexibilität, ein ggf. höheres Maß an Know-how, gute Infrastruktur und externe Praxissicht.[132] Dies kann auch als Zwischenschritt zum Aufbau eines rein staatlichen Automatisierungsgremiums gesehen werden, indem durch die anhaltende Vermittlung relevanter (Er)Kenntnisse nach und nach Expertenwissen an den Staat fließt, bis dieser einen hinreichenden eigenen Unterbau an Expertenwissen aufweist, um die rein staatliche Variante zu implementieren.[133] In diesem Modell ließe sich auch nach unterschiedlichen Stufen der strukturellen Automatisierungsentscheidungen unterscheiden. So ist beispielsweise im gegenwärtigen MAKI-Projekt vorgesehen, dass zwar strukturelle Entwicklung, Betrieb und Überwachung einerseits durch den interdisziplinären Verbund der Kooperationspartner erfolgt. Zur stärkeren Realisierung der individuellen richterlichen Unabhängigkeit der die KI einsetzenden Richterinnen und Richter findet ein Training des Systems vom erkennenden Richter bzw. von der erkennenden Richterin statt (über eine sog. „One shot“-Annotation, d. h. das einmalige Trainieren des Systems anhand einer Musterakte wird das System trainiert, damit es in Folgeverfahren eine eigenständige Zuordnung zu einer Verfahrenskategorie vornehmen kann). Für den einzelnen Richter bzw. die einzelne Richterin ist dieses System allerdings mit höherem Aufwand verbunden als die zentrale Gestaltung eines entsprechenden Unterstützungssystems durch eine „Automatisierungskammer“ bzw. ein interdisziplinäres Expertengremium.

(ii) Handlungsempfehlung

Ein **richterliches Governance-Modell** für die Aufsicht und Kontrolle eines informationstechnischen Systems zu etablieren, **ist unbedingt anzuraten**, da es sich bei der Umsetzung richterlicher Mitwirkung, Kontrolle und Aufsicht um eine der zentralen roten Linien handelt, die unbedingt Beachtung finden muss.
Welches der genannten Governance-Modelle in welcher konkreten Ausgestaltungsvariante das geeignetste ist, kann nicht pauschal festgehalten werden, da eine den rechtlichen Voraussetzungen entsprechende effektive Einbindung von Richterinnen und Richtern in die Beschaffung, die Entwicklung, den Betrieb und die Überwachung des informationstechnischen Systems sehr stark davon abhängig ist, für welche konkreten Funktionen und Einsatzszenarien das System gestaltet wird und welche Funktionsweise für das System letztlich charakteristisch ist (regelbasiertes System vs. System maschinellen Lernens vs. Mischsystem). In unserem Abschlussbericht des ersten Arbeitspakets (Milestone 1) des Forschungsprojekts „Künstliche Intelligenz und richterliche Entscheidungsfindung“ ist eingehend herausgearbeitet worden, dass sich die Einsatzgrenzen und die Anforderungen an den

[131] Vgl. *Wolff*, Algorithmen als Richter, 1. Aufl. 2022, S. 182f.
[132] *Spoenle*, OLGA – der KI-Assistent für Dieselverfahren, in: Reuß/Laß (Hrsg.), Göttinger Kolloquien zur Digitalisierung des Zivilverfahrensrechts Kolloquien im Sommersemester 2022 | Kolloquien im Wintersemester 2022/2023, Göttingen 2024, S. 175.
[133] Vgl. *Wischmeyer*, AöR 143 (2018), 1, 62f.

legalen Betrieb eines informationstechnischen Systems v. a. anhand dreier Faktoren bestimmen:[134] erstens, nach dem **funktionalen Charakter der Tätigkeit** (rein administrative oder inhaltliche Funktionalität), zweitens, nach dem **Grad der Automatisierung,** sowie drittens, nach der **Komplexität des jeweiligen Anwendungsfeldes**. Je stärker das System sachlich-inhaltliche Aufgaben wahrnimmt, je höher der Grad der Automatisierung ist und je komplexer das Anwendungsfeld gestaltet ist, in dem ein System Funktionen erfüllen soll, desto höher sind die Anforderungen zur Einhaltung der ethischen Einsatzgrenzen und der rechtlichen Rahmenbedingungen. Dies bedeutet auch, dass die Anforderungen an die Gewährleistung richterlicher Aufsicht und Kontrolle immer wichtiger werden, je stärker das System sachlich-inhaltliche Aufgaben wahrnimmt, je stärker automatisiert die Aufgabenerledigung erfolgt und je komplexer das Einsatzszenario gestaltet wird.
So ist beispielsweise für einen entscheidungs*ersetzenden* Einsatz, der nach der hier vertretenen Auffassung derzeit ohnehin lediglich in ganz einfach gelagerten Fallkonstellationen mittels regelbasierter Systeme denkbar ist (z. B. teilweise im Falle von Fluggastrechteverfahren),[135] durchaus vorstellbar, Richterinnen und Richter als Experten gemeinsam mit einem interdisziplinären Entwicklungsteam an der Gestaltung des Systems im Rahmen einer Entwicklungsarbeitsgruppe zu beteiligen (idealerweise bundeseinheitlich, wenn bundesweit dieselben Rechts- und Tatsachenfragen aufgeworfen werden). Die v. a. inhaltliche Überwachung des Systems könnte dann jedoch einzelnen, für die Aufgaben im Einklang mit dem Recht auf einen gesetzlichen Richter, Art. 101 Abs. 1 S. 2 GG, in den Geschäftsverteilungsplänen der dezentralen Gerichtsorganisation zuvor bestimmten Richterinnen und Richtern zugewiesen werden (ggf. auch hier zentriert bei einem speziellen Gericht/Bundesland; siehe hierzu auch weitergehend unter C.II.6). Über ein regelmäßiges Monitoring und ein Feedbacksystem mit den das System technisch und inhaltlich überwachenden Personen ließe sich in dem interdisziplinären Entwicklungs- und Überwachungsteam eine regelmäßige Prüfung und Anpassung des Systems vorsehen und gestalten, z. B. auch bei Rechtsänderungen oder technischem Anpassungsbedarf. Eine Notwendigkeit für sog. „Automatisierungskammern“ besteht in diesen Fällen nicht. Wegen der tatsächlichen und der rechtlichen Überschaubarkeit des Einsatzfeldes wären hier mehrere Entwicklungs- und Überwachungsgremien eher hinderlich und würden unnötige staatliche Ressourcen verbrauchen. Hiervon ist in diesen Fällen abzuraten.
Demgegenüber kommt ein entscheidungs*unterstützender* Einsatz auch mit Systemen maschinellen Lernens und in vielfachen Einsatzszenarien mit unterschiedlichen Graden an Komplexität, Funktionalität und Automatisierung in Betracht. Es ergibt sich eine Bandbreite an Einsatzszenarien und anhand der von den genannten Faktoren abhängigen Einsatzmatrix ergeben sich auch unterschiedlich geeignete Governance-Modelle. Handelt es sich beispielsweise um ein System, das schlicht administrative Aufgaben ausübt, z. B. ein agentischer KI-Verfahrensassistent, der automatisiert eingehende Verfahrensinformationen benennt (d. h. Metadaten zuweist), die Information dem relevanten Verfahren zuweist und in der elektronischen Akte an der relevanten Stelle abrufbar speichert, dürfte die Überlassung der inhaltlichen Überwachung an den jeweiligen erkennenden Richter bzw. die erkennende Richterin genügen, denn Fehlzuweisungen, fehlgehende Speicherungen etc. werden im

[134] Siehe hierzu ausführlich *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 12.
[135] Vgl. *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 88, 107, 145, 171ff., 219f., 223, 339, 357.

laufenden Betrieb problemlos erkannt und korrigiert werden können.[136] Die richterliche Einbindung im Entwicklungsprozess (i. R. e. interdisziplinären Entwicklungsarbeitsgruppe) dürfte in diesem Fall zwar mangels Berührung des Kernbereichs der Rechtsprechung rechtlich nicht zwingend geboten, gleichwohl dennoch zu empfehlen sein, aufgrund der rechtlichen Expertise, um die Verfahrensabläufe zu gestalten. Auch hier ist eine „Automatisierungskammer“ o. ä. nicht erforderlich und würde unnötigen Ressourcenaufwand bedeuten. Sind allerdings auf der anderen Seite der Bandbreite an denkbaren Funktionalitäten Systeme in Rede, die vollautomatisiert eingehende Akten analysieren, in Verfahrenskategorien zuordnen und konkrete, auf den jeweiligen Fall vorgesehene Entscheidungsvorschläge unterbreiten, sind die Anforderungen an eine richterliche Aufsicht und Kontrolle deutlich höher. Hier kann sich insbesondere aufgrund der komplexeren Anwendungsszenarien ein dauerhaftes Automatisierungsgremium, wie die vorgeschlagenen „Automatisierungskammern“, durchaus anbieten, die auch dezentral organisiert sein können, insbesondere dann, wenn auch regionale Besonderheiten bei der Entscheidungsfindung eine Rolle spielen (z. B. bestimmte Lebenshaltungskosten an einem Ort bzw. ein relevanter Mietspiegel).
Die adaptierte – und nicht pauschale – Ausgestaltung von Aufsicht und Kontrolle, insbesondere (auch) anhand der Faktoren der dargestellten dreigliedrigen Funktionsmatrix, die im Rahmen des ersten Milestones entwickelt wurde, im Sinne eines mehrschichtigen bzw. Mehr-Ebenen-Konzepts entspricht bzw. resoniert auch den Erkenntnissen aus der psychologischen bzw. interdisziplinären Forschung.[137] Die psychologische bzw. interdisziplinäre Forschung zum Thema der Aufsicht und Kontrolle über KI-Systeme hat zudem verschiedene Faktoren identifiziert, deren Beachtung für die wirksame Wahrnehmung der Aufgaben zentral ist, namentlich menschliche Faktoren (z. B. Wissen und Schulungen; siehe hierzu insbesondere Abschnitt C.II.2.a.bb(3)), technische Faktoren (z. B. Herstellung von Transparenz hinsichtlich des KI-Systems, Systemschnittstellen für die Aufsichtspersonen; siehe hierzu insbesondere die Abschnitte C.II.2.c.bb(1) und C.II.4) und organisatorische Faktoren (z. B. verfügbare Zeit und Ressourcen; siehe hierzu insbesondere Abschnitt C.II.9), deren Gehalt als wesentliche Aspekte auch bereits im Rahmen des ersten Milestones des vorliegenden Forschungsprojekts in den Blick genommen wurden und im vorliegenden Milestone weiter betrachtet werden.[138]

---

[136] Zu Vorteilen und Risiken im Zusammenhang mit KI-Agenten, der Begrifflichkeit von KI-Agenten, einer Betrachtung von KI-Assistenten in der öffentlichen Verwaltung und insbesondere zur Thematik menschlicher Aufsicht nach der KI-VO in Bezug auf agentische KI, *Meinel/Baum/Lauber-Rönsberg/Stenzel/Langer*, Human Oversight of AI Agents under the EU AI Act, 2025, abrufbar unter https://dx.doi.org/10.2139/ssrn.7500959 (Abruf am 21.10.2026, 22:52 Uhr).

[137] Zu einem interdisziplinär entwickelten Konzept für wirksame Aufsicht und Kontrolle auch m. w. H. auf Risiken und Problemen mit Aufsicht und Kontrolle *Gaube/Langer/Miller/Baum/Dachselt/Feit/Gadiraju/Kaur/Keane/Landers/Laux/Liao/Lim/Onnasch/Schrills/Sonenberg/Tan/Tintarev/Xiao/Zhang*, Keeping an Eye on AI: A Framework for Effective Human Oversight of AI Systems, 2026, abrufbar unter https://doi.org/10.48550/arXiv.2605.16278 (Abruf am 24.10.2026, 23:00 Uhr); zu einer interdisziplinär angelegten Definition zur menschlichen Aufsicht siehe a. a. O. S. 2f.; zur mehrschichtigen bzw. mehr-Ebenen-Ausgestaltung a. a. O. S. 4ff.

[138] *Gaube/Langer/Miller/Baum/Dachselt/Feit/Gadiraju/Kaur/Keane/Landers/Laux/Liao/Lim/Onnasch/Schrills/Sonenberg/Tan/Tintarev/Xiao/Zhang*, Keeping an Eye on AI: A Framework for Effective Human Oversight of AI Systems, 2026, abrufbar unter https://doi.org/10.48550/arXiv.2605.16278 (Abruf am 24.10.2026, 23:00 Uhr), S. 3, 8f.; siehe zudem zu Faktoren die eine wirksame menschliche Aufsicht bedingen *Sterz/Baum/Biewer/Hermanns/Lauber-Rönsberg/Meinel/Langer*, On the Quest for Effectiveness in Human Oversight: Interdisciplinary Perspectives, 2024, abrufbar unter https://doi.org/10.48550/arXiv.2404.04059 (Abruf am 22.10.2026, 20:38 Uhr), S. 3ff.; Siehe zu den Bedingungen wirksamer menschlicher Aufsicht auch Abschnitt C.I.4.d.bb.

Damit lässt sich festhalten, dass ein einheitliches richterliches Governance-Modell, das alle Einsatzszenarien angemessen abdecken könnte, nicht existiert. Es ist daher **ein auf den konkret geplanten Einsatz abgestimmtes Modell zu entwickeln und zu etablieren**. Aus dem hier interessierenden Blickwinkel muss ein solches Modell sicherstellen, dass richterliche Aufsicht und Kontrolle im Rahmen der Beschaffung, des Betriebs, der Überwachung sowie erforderlichenfalls der Entwicklung des konkreten Systems effektiv sichergestellt sind.

(2) Ansiedelung der Entwicklung und des Betriebs des informationstechnischen Systems

Weitere Handlungsoptionen sind für die grundlegende **Ansiedelung der Entwicklung und des Betriebs des informationstechnischen Systems** in den Blick zu nehmen.

(i) Handlungsoptionen

Auch hierzu bestehen vielfältige Möglichkeiten, die teils in der Literatur bereits diskutiert werden:
Einerseits ist denkbar, *Entwicklung und Betrieb bei der Exekutive bzw. bei privaten Dritten* zu verorten.[139] Hierbei ist jedoch zu beachten, dass vor dem Hintergrund des Gewaltenteilungsgrundsatzes sowie insbesondere der sachlichen (und persönlichen) richterlichen Unabhängigkeit nach Art. 97 GG strenge Anforderungen gelten.[140] So urteilte der Bundesgerichtshof zur Überlassung eines EDV-Netzes durch die Verwaltung an die Justiz, dass mit dieser keine unzulässige inhaltliche Kontrolle richterlicher Dokumente einhergehen darf und die Überlassung „dem sachgerechten Betrieb und der ordnungsgemäßen Verwaltung des EDV-Netzes“ dienen muss.[141] Bereits, „wenn ein besonnener Richter durch ein Gefühl des unkontrollierbaren Beobachtetwerdens (…) von der Verwendung der ihm zur Erfüllung seiner richterlichen Aufgaben zur Verfügung gestellten Arbeitsmittel abgehalten würde“, kann die sachliche Unabhängigkeit gem. Art. 97 Abs. 1 GG verletzt sein.[142] Freilich liegen die notwendigen Kapazitäten für die Entwicklung und den Betrieb informationstechnischer Systeme indes regelmäßig eher bei der Exekutive bzw. Dritten als bei der Judikative selbst.[143] So bedarf es beispielsweise der Programmierung komplexer sowie umfassender Expertensysteme, um bereits kleine Rechtsfragen zu beantworten.[144] Wird das IT-System der Verwaltung überlassen, sind klare Regelungen behördlicher Zugriffsrechte auf das System samt weitreichender Auflagen (u. a. inhaltlicher Zugriff Dritter auf richterliche Dokumente nur bei Betriebsnotwendigkeit für das EDV-Netz

[139] Vor dem Hintergrund u. a. der finanziellen Abhängigkeit gegenüber öffentlich-privaten Partnerschaften im Rahmen der Beschaffung von KI zur besonderen Vorsicht mahnend *Turbék*, Justiz und KI, Friedrich-Naumann-Stiftung für die Freiheit (Hrsg.), 2026, abrufbar unter https://shop.freiheit.org/#!/Publikation/2184 (Abruf am 09.10.2026, 15:40 Uhr), S. 31.
[140] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 131ff., 188ff., 195ff.
[141] *BGH*, MMR 2012, 128, 128.
[142] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 131ff., 188ff.
[143] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 131ff., 195ff.; *Horn*, ZZP 136 (2023), 459, 478; siehe auch *Stürner*, GPR 2025, 62, 67 mit Hinweis auch auf die Belastungen infolge der KI-Verordnung.
[144] Kaulartz/Braegelmann AI und Machine Learning-HdB/*Rühl*, 1. Aufl. 2020, Kap. 14.1 Rn. 13.

sowie Untersagung der Weitergabe richterlicher Dokumente) zu empfehlen.[145] Diese Regelungen sind schriftlich niederzulegen und ihre Einhaltung ist mit Beteiligung (etwa von gewählten Vertretern) der Richterschaft zu überwachen.[146] Teilweise wird diesbezüglich vertreten, dass der Verwaltung lediglich der Zugriff auf äußere Daten gestattet werden sollte, ohne ihr eine inhaltliche Einsicht zu gewähren.[147] Jedenfalls sind der Judikative bei einer Systemüberlassung an die Verwaltung effektive Kontroll- und Mitsprachemöglichkeiten einzuräumen.[148] Von maßgeblichem Vorteil kann hierbei eine wirksame Absicherung justiziellen Einflusses sein, etwa durch entsprechende vertragliche Gestaltungen und technische Lastenhefte mit Auditierbarkeit,[149] Protokollierung, dokumentierte Testverfahren und -läufe[150] sowie Freigabeverfahren.[151] Die Forderung, dass die Auswahl und Beschaffung der Systeme stets von den Gerichten selbst vorzunehmen sei, ist zu weitgehend, vielmehr genügt die Einhaltung der o. g. roten Linien, wobei eine hinreichende, maßgebende sowie mitprägende Involvierung von Richterinnen und Richtern in Projektgruppen für die Auswahl sowie Beschaffung der Software ausreichend ist.[152]

Aufgrund der zunehmenden Komplexität und Zentralisierung von IT-Anwendungen kann eine praktische Wahrnehmung der Kontrollbefugnisse indes auf Schwierigkeiten stoßen, weshalb das Einsetzen von mit IT-Angelegenheiten schwerpunktmäßig befassten **richterlichen IT-Beauftragten** mit Beratungs-, Kontroll- und Überwachungsfunktion vorzugswürdig erscheint.[153] Dies kann insbesondere (auch) durch (insbesondere interdisziplinär besetzte) Automatisierungskammern oder Automatisierungsinstitutionen (s. o.) erfolgen. Die Vorzugswürdigkeit eines derartigen Einsatzes gilt mit Blick auf die Komplexität von Systemen künstlicher Intelligenz und die damit einhergehenden, dynamischen sowie neuen Fragestellungen umso mehr; die praktischen Schwierigkeiten bei der Wahrnehmung richterlicher Kontrollbefugnisse sind entsprechend erheblicher im Vergleich zur Lage bei herkömmlichen IT-Anwendungen.[154]

Sollte eine hinreichende Aufsicht und Kontrolle nicht sicherzustellen sein, scheidet eine Nutzung der Systeme aus oder sie sind in den judikativen Rahmen einzugliedern.[155] Die für den Einsatz der informationstechnischen Systeme notwendige Kontrolle und Aufsicht sollte daher nicht stiefmütterlich behandelt werden.

---

[145] Vgl. *BGH*, MMR 2012, 128, 128; *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 131ff., 188ff., 195ff.

[146] Hornung/Schallbruch/*Müller*, IT-Sicherheitsrecht, 2. Aufl. 2024, § 27 A. I. 2. Rn. 21, 27; *BGH,* MMR 2012, 128, 129.

[147] *v. Notz*, JZ 2017, 607, 613.

[148] *Heckmann*, in: Bräutigam (Hrsg.), IT-Outsourcing und Cloud-Computing, Teil 10 Rn. 167 m. w. N.

[149] Vgl. *Martini*, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz, 1. Aufl. 2019, S. 323.

[150] Vgl. *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 373.

[151] Vgl. auch Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 34f.; Abgestimmte Standards können auch hierbei von Vorteil sein, vgl. auch *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 21f.

[152] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 133.

[153] Hornung/Schallbruch/*Müller*, IT-Sicherheitsrecht, 2. Aufl. 2024, § 27 A. I. 2. Rn. 13f., 23.

[154] Vgl. in diesem Zusammenhang auch Bomhard/Pieper/Wende/*Bomhard*/*Wietzke*, 1. Aufl. 2025, KI-VO Art. 14 Rn. 5ff.

[155] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 132f.

Ihr sollte ein entsprechender Stellenwert in der Justizorganisation eingeräumt werden, etwa indem richterliches Personal nach und in Verbindung mit regelmäßigen entsprechenden Fortbildungsmaßnahmen für diese Tätigkeit freigestellt wird sowie ggf. in Verbindung mit entsprechenden besoldungstechnischen Zuschlägen. Das Einräumen des entsprechend hohen Stellenwerts einer adäquaten Aufsicht steht insbesondere auch mit den Wertungen der KI-VO in Einklang. Bislang war nach Art. 4 KI-VO die Vermittlung der nötigen Kenntnisse an ihr Personal für Anbieter und Betreiber mit Blick auf alle (nicht nach Art. 2 KI-VO dem Anwendungsbereich der KI-VO entzogenen) KI-Systeme i. S. d. Art. 3 Abs. 1 KI-VO verpflichtend.[156] Zwar wurde infolge der Neufassung durch die Digital-Omnibus-Verordnung zur KI[157] diese Verpflichtung derart abgeschwächt, dass nur noch Maßnahmen zu ergreifen sind, um die Entwicklung der KI-Kompetenz zu unterstützen, unter Klarstellung in Art. 4 Abs. 1 S. 2 KI-VO, dass keine Verpflichtung für die Garantie einer bestimmten KI-Kompetenz besteht. Dies erfolgte jedoch nach ErwGr. 8 S. 3, 4 der Digital-Omnibus-Verordnung zur KI, da eine strenge Verpflichtung, ein ausreichendes Maß an KI-Kompetenz sicherzustellen, nicht für alle Arten von Anbietern und Betreibern von KI geeignet erschien und aufgrund des zusätzlichen Befolgungsaufwands insbesondere für kleinere Unternehmen. Wertungstechnisch bleibt daher der vorliegende Bereich unberührt, zumal ErwGr. 8 S. 2 der Digital-Omnibus-Verordnung zur KI ausdrücklich (weiterhin) bekräftigt, dass der Erwerb von KI-Kompetenz angefangen bei der allgemeinen und beruflichen Bildung bis hin zum lebenslangen Lernen, unerlässlich ist, um Anbieter, Betreiber und andere betroffene Personen mit den notwendigen Fähigkeiten auszustatten, damit sie fundierte Entscheidungen über den Einsatz von KI-Systemen treffen können. Zudem ist auch (weiterhin) (im Hochrisiko-KI-Bereich) der Grundsatz der menschlichen Aufsicht in Art. 14 KI-VO niedergelegt und durch Art. 26 Abs. 2 KI-VO zur konkreten Handlungs- und Überwachungspflicht des Betreibers ausgeformt.[158] Der erhebliche Stellenwert funktionierender menschlicher Aufsicht wird hierbei durch die Erwägungsgründe der KI-VO, insbesondere ErwGr. 27 S. 3–5, S. 11–13, ErwGr. 66, ErwGr. 73, ErwGr. 91 sowie maßgeblich ErwGr. 1 der KI-VO, unterstrichen, indem die zentrale Zielsetzung der menschenzentrierten KI maßgeblich durch die menschliche Aufsicht erreicht werden soll.[159]

Als weitere Handlungsoption kommt die *originäre Ansiedelung der Entwicklung und des Betriebs der informationstechnischen Systeme bei der Judikative* in Betracht.[160] Hierbei kann auch ein abgestufter Ansatz verfolgt werden und jedenfalls Systeme mit Funktionen im entscheidungsersetzenden sowie im inhaltsbezogenen, entscheidungsunterstützenden Bereich weitestgehend bis hin zu ausschließlich in der Hand der Judikative verortet sein.

---

[156] Martini/Wendehorst/*Wendehorst*, 2. Aufl. 2026, KI-VO Art. 4 Rn. 19ff.; HK-KI-VO/Hilgendorf/Härtlein/*Hilgendorf*/*Härtlein*, 1. Aufl. 2025, KI-VO Art. 4 Rn. 1.

[157] VERORDNUNG (EU) 2026/1744 DES EUROPÄISCHEN PARLAMENTS UND DES RATES vom 8. Juli 2026 zur Änderung der Verordnungen (EU) 2024/1689, (EU) 2018/1139 und (EU) 2023/1230 im Hinblick auf die Vereinfachung der Umsetzung harmonisierter Vorschriften für künstliche Intelligenz (Digital-Omnibus-Verordnung zur KI), ABl. L, 2026/1744, 24.7.2026.

[158] BeckOK KI-Recht/*Denga*, 7. Ed. 1.8.2026, KI-VO Art. 26 Rn. 32.

[159] BeckOK KI-Recht/*Buchner*, 7. Ed. 1.8.2026, KI-VO Art. 14 Rn. 8, 12ff.

[160] *Radke*, jM 2016, 8, 12; *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 195ff.

(ii) Handlungsempfehlung

Erneut kann hier auf die Funktionsmatrix anhand des funktionalen Charakters der Tätigkeit, des Grades der Automatisierung sowie der Komplexität des jeweiligen Anwendungsfeldes des ersten Milestone-Berichts des MAKI-Forschungsprojekts Bezug genommen werden.[161] Mit steigendem Funktionsgrad steigt auch die Vorzugswürdigkeit, dass die informationstechnischen Systeme vollständig in der Hand der Judikative liegen.[162] Zwar kann verzeichnet werden, dass das Outsourcing der IT-Infrastruktur bislang wohl keine nachhaltigen Vorbehalte in der Praxis im Hinblick auf die richterliche Unabhängigkeit aufgeworfen hat.[163] Auch ist zuzugeben, dass eine vollständige Sicherheit vor externem Zugriff auf die richterliche Tätigkeit wohl noch nie bestand.[164] Jedoch werden schon bislang durch den Einsatz moderner Technik Arbeitsinhalte sowie Arbeitsergebnisse zunehmend vorgeprägt.[165] Mit dem Einsatz künstlicher Intelligenz in der richterlichen Entscheidungsfindung wird diese Vorprägung präsumtiv noch einmal weiter erheblich intensiviert. Auch kann bereits durch eine einzige unbefugte Einflussnahme auf die Systeme eine erhebliche Folgewirkung gezeitigt werden.[166] Durch eine Einflussnahme auf ein System können bei zunehmendem Einsatz von KI, auch in Verbindung mit *Bias*, ganze Zweige der Justiz unbefugt beeinflusst werden. Beispielhaft könnte hierfür etwa ein landes- oder bundesweit im Einsatz befindliches System angeführt werden, das im Rahmen der Strafzumessung eingesetzt ist und dahingehend manipuliert wird, bei Vorliegen bestimmter Namen pauschal einen höheren oder niedrigeren Strafrahmen auszugeben. Im Rahmen des Zivilprozesses könnte z. B. eine Manipulation vorgenommen werden, dass bei einer Parteienstellung von bestimmten (juristischen oder natürlichen) Personen, das System diese wo irgend möglich im Rahmen der Entscheidungsfindung (z. B. durch eine unvorteilhafte Darstellung des Vortrags der Gegenseite bei der Ausgabe einer Relationstabelle oder Bejahung bzw. Verneinung des Vorliegens von Einwendungen oder Einreden) sowie im Rahmen der Entscheidung selbst bevorzugen soll. Mit Blick auf die strengen Grenzen der richterlichen Unabhängigkeit erscheint es (somit) auch hier vorzugswürdig, strenge Maßstäbe anzulegen. Diese Strenge ist eine Implikation der verfassungsrechtlich garantierten richterlichen Unabhängigkeit in Verbindung mit den praktischen Schwierigkeiten bei der Kontrolle informationstechnischer Systeme und sollte nicht als Misstrauen gegenüber der Exekutive oder privaten Anbietern missverstanden werden. Ein solcher „Königsweg“, der Ansiedelung bei der Judikative, wird bislang z. B. mithilfe des beim Oberlandesgericht Köln angesiedelten Zentralen IT-Dienstleisters für die Justiz des Landes Nordrhein-Westfalen (ITD)[167] (für fast alle IT-Aufgaben)

[161] Zusammenfassend siehe *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 398.
[162] Vgl. in diesem Zusammenhang auch *Dötterl*, RDi 2026, 414, 416f., der die Berührung des Kernbereichs der richterlichen Unabhängigkeit in der Intensität als entsprechend steigend ansieht, je weiter die inhaltliche Vorbewertung durch das KI-System erfolgt.
[163] *Radke*, jM 2016, 8, 10.
[164] *Radke*, jM 2016, 8, 9ff.
[165] *Heckmann*, in: Bräutigam (Hrsg.), IT-Outsourcing und Cloud-Computing, Teil 10 Rn. 158; siehe in diesem Zusammenhang zur Beeinflussung des Entscheidungsvorgangs auch *Dötterl*, RDi 2026, 414, 415.
[166] Vgl. *Turbék*, Justiz und KI, Friedrich-Naumann-Stiftung für die Freiheit (Hrsg.), 2026, abrufbar unter https://shop.freiheit.org/#!/Publikation/2184 (Abruf am 09.10.2026, 15:40 Uhr), S. 17.
[167] Vgl. auch Internetauftritt des ITD, Organisation, abrufbar unter https://www.justiz.nrw.de/Gerichte_Behoerden/zentraler_dienstleister/ah20_organisation (Abruf am 28.01.2026).

beschritten.[168] Auch das Land Niedersachsen verfügt über einen bei der Justiz angesiedelten Zentralen IT-Betrieb (ZIB).[169] Diese rechtliche Analyse stützt auch ein Blick auf die rechtsethischen Prinzipien. Das **Prinzip der technologischen Souveränität**,[170] d. h. die souveräne, d. h. eigenständige und von Externen möglichst unbeeinflusste Entscheidung über den Einsatz und den Betrieb informationstechnischer Systeme wird durch einen justizeigenen Betrieb, der weitgehend unabhängig von (ausländischen) privaten bzw. staatlichen Akteuren ist, stärker verwirklicht als durch einen cloudbasiertern Betrieb bei einem (ausländischen) privaten Akteur.[171] Die Einfluss- und Steuerungsmöglichkeiten sind bei einem justizeigenen Betrieb erheblich höher. **Je stärker das System inhaltsbezogene Aufgaben wahrnimmt, desto näher rückt die vollständige Ansiedelung des informationstechnischen Systems in der judikativen Hand als vorzugswürdiger „Königsweg“.**

(3) Überwachung der Arbeitsergebnisse; Umgang mit Bias und sonstigen Fehlermustern

Mit Blick auf die Überwachung der systemseitigen Arbeitsergebnisse sowie im Umgang mit Bias und sonstigen Fehlermustern sind folgende **Handlungsempfehlungen** zu geben:
Es sollten dringend die für die Überwachung und Kontrolle **nötigen Kenntnisse vermittelt** werden.[172] Hierzu zählen einerseits die fachlichen Kenntnisse der jeweiligen Personen im relevanten Anwendungsfeld (z. B. die Rechtskenntnisse von Richterinnen und Richtern), aber v. a. auch die Kenntnisse zur Funktionsweise informationstechnischer Systeme, deren Stärken und deren Schwächen.[173] Nur wer über fundierte Kenntnisse des rechtlichen Anwendungsfeldes und über Kenntnisse der technischen Funktionsweise der eingesetzten Systeme verfügt, wird in der Lage sein, die erforderlichen Überwachungsaufgaben ausführen zu können. Die Vermittlung der nötigen Kenntnisse ist hierbei auch im Rahmen der europäischen KI-VO vorgesehen. Art. 4 KI-VO gilt für alle (nicht nach Art. 2 KI-VO dem Anwendungsbereich der KI-VO entzogenen) KI-Systeme i. S. d. Art. 3 Abs. 1 KI-VO und verpflichtet gem. Art. 4 Abs. 1 S. 1 KI-VO Anbieter sowie Betreiber zum Ergreifen von Maßnahmen, um die Entwicklung der KI-Kompetenz ihres Personals und anderer Personen zu unterstützen.[174] Im Bereich der Hochrisiko-KI-Systeme wird diese Norm insbesondere ergänzt um das Zusammenspiel aus Art. 14 (insb. Abs. 1 und Abs. 4) KI-VO, der die Möglichkeit der menschlichen Aufsicht als Anforderung an das KI-System selbst vorschreibt und damit sichert, sowie Art. 26 Abs. 2 KI-VO, der Betreiber verpflichtet, nur entsprechend kompetenten natürlichen Personen die Aufsicht über die Hochrisiko-KI-Systeme zu übertragen und diesen die erforderliche Unterstützung zukommen zu

[168] *Radke*, jM 2016, 8, 12.
[169] https://zib.niedersachsen.de/startseite/service_neu/impressum-197077.html (Abruf am 06.07.2026).
[170] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 69.
[171] Vgl. zu der Thematik auch das Bund/Länder-Vorgehen *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 22ff.
[172] Siehe auch die Schulung von Aufsichtspersonen im Rahmen von wirksamer Aufsicht als entscheidend bezeichnend *Sterz/Baum/Biewer/Hermanns/Lauber-Rönsberg/Meinel/Langer*, On the Quest for Effectiveness in Human Oversight: Interdisciplinary Perspectives, 2024, abrufbar unter https://doi.org/10.48550/arXiv.2404.04059 (Abruf am 22.10.2026, 20:38 Uhr), S. 7.
[173] Siehe auch *Köhler*, Richterliche Verantwortung und Künstliche Intelligenz, 1. Aufl. 2026, S. 90f., 205.
[174] Martini/Wendehorst/*Wendehorst*, 2. Aufl. 2026, KI-VO Art. 4 Rn. 19ff.

lassen.[175] Inhaltlich sollten zudem, neben den technischen sowie rechtlichen Kompetenzen (insbesondere auch Kenntnisvermittlung bezüglich KI-adressierender Regelungen, wie die KI-VO umfassend), auch weitere thematisch involvierte sowie angrenzende Disziplinen bei der Kenntnisvermittlung berücksichtigt werden und neben den technischen und rechtlichen Fähigkeiten auch Wert auf die Vermittlung von sozialen und ethischen Implikationen gelegt werden.[176] Des Weiteren sollte darauf geachtet werden, dass insbesondere im Fall der entscheidungsunterstützenden Funktionen kein langfristiger Kompetenzverlust durch die Auslagerung der Wahrnehmung von Aufgaben an informationstechnische Systeme eintritt (sog. „Deskilling“) bzw. Gegenmaßnahmen ergriffen werden; der Kompetenzverlust kann sich auch wiederum negativ auf die Wahrnehmung der Überwachungsaufgaben auswirken.[177]

Bei der Bestimmung **geeigneter Schulungsmaßnahmen** sind die technischen Fachkenntnisse, der Erfahrungsstand und die Aus- sowie Fortbildung des zu schulenden Personenkreises sowie der Kontext des KI-Systemeinsatzes maßgeblich einzubeziehen; die Schulungsmaßnahmen sind abhängig vom zu schulenden Personenkreis zu wählen und anzupassen.[178] Die involvierten Kreise, d. h. rechtswissenschaftlich ausgebildete Richterinnen und Richter, aber auch beispielsweise das entsprechend notwendige Technik-Personal (sowie jeweils zudem das Führungspersonal) sind im Umgang mit den informationstechnischen Systemen, abhängig vom jeweiligen Bedarf zu schulen – eine individuelle Bedarfsanalyse sowie

---

[175] HK-KI-VO/Hilgendorf/Härtlein/*Hilgendorf/Härtlein*, 1. Aufl. 2025, KI-VO Art. 14 Rn. 1, Art. 26 Rn. 3ff.; Zweifel an der Wirksamkeit darstellend *Dötterl*, RDi 2026, 414, 418 m. w. H.; die Umstrittenheit hinsichtlich der Wirksamkeit von Maßnahmen menschlicher Aufsicht aufzeigend sowie ein dynamisches Aufsichtsmodell zur Ausfüllung der menschlichen Aufsicht nach der KI-VO anhand eines Lebenszyklusansatzes, der antizipierende, betriebsbezogene sowie evaluierende Funktionen umfasst, darstellend *Meinel/Baum/Lauber-Rönsberg/Stenzel/Langer*, Human Oversight of AI Agents under the EU AI Act, 2025, abrufbar unter https://dx.doi.org/10.2139/ssrn.7500959 (Abruf am 21.10.2026, 22:52 Uhr), S. 10ff. sowie *Sterz/Baum/Biewer/Hermanns/Lauber-Rönsberg/Meinel/Langer*, On the Quest for Effectiveness in Human Oversight: Interdisciplinary Perspectives, 2024, abrufbar unter https://doi.org/10.48550/arXiv.2404.04059 (Abruf am 22.10.2026, 20:38 Uhr), S. 3 (die zudem die Nutzung der von ihnen geschaffenen Aufsicht-Wirksamkeitsdefinition für Art. 14 KI-VO nahelegen, S. 9f.).

[176] Im unternehmerischen Kontext anführend *Hofer/Kirchmair*, ZfPC 2025, 270, 273; Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 24.

[177] *Dötterl*, RDi 2026, 414, 418; *Gaube/Langer/Miller/Baum/Dachselt/Feit/Gadiraju/Kaur/Keane/Landers/Laux/Liao/Lim/Onnasch/Schrills/Sonenberg/Tan/Tintarev/Xiao/Zhang*, Keeping an Eye on AI: A Framework for Effective Human Oversight of AI Systems, 2026, abrufbar unter https://doi.org/10.48550/arXiv.2605.16278 (Abruf am 24.10.2026, 23:00 Uhr), S. 12; zur Begrifflichkeit des „Deskilling“ siehe *Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen*, Deskilling, abrufbar unter https://www.lernen-digital.nrw/glossary-term/100 (Abruf am 24.09.2026, 15:42 Uhr).

[178] *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 181; an dieser Stelle sei zudem Frau Isabelle Biallaß für ihren Vortrag „AI-Literacy in der Justiz – Anforderungen an Aus- und Fortbildung“ bei der AI@Law-Conference vom 29.01.2026 in München gedankt, in dem sie insbesondere über die Praxiserfahrungen bezüglich der KI-Kompetenz im richterlichen Kontext wertvolle Einblicke gab; siehe auch *Hofer/Kirchmair*, ZfPC 2025, 270, 273; die Förderung der KI-Kompetenz, bedarfsgerechte Fortbildungs- und Sensibilisierungsangebote umfassend, ist auch als wichtig identifiziert worden von *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 34f.

ein individueller Schulungszuschnitt sind notwendig.[179] Mit Blick auf die präsumtiv nachhaltende Relevanz informationstechnischer Systeme ist ein frühzeitiger Schulungs-Ansatz sinnvoll, bereits im Rahmen der Ausbildung bzw. im Rahmen des Studiums entsprechende Kompetenzen vermittelt zu bekommen.[180] Entsprechendes wird auch im Zusammenhang mit der KI-VO ausdrücklich genannt. In ErwGr. 8 S. 2 der Digital-Omnibus-Verordnung zur KI wird der Erwerb von KI-Kompetenz, angefangen bei der allgemeinen und beruflichen Bildung bis hin zum lebenslangen Lernen, als unerlässlich für die Ausstattung mit den notwendigen Fähigkeiten bezeichnet, damit fundierte Entscheidungen über den Einsatz von KI-Systemen getroffen werden können. In der (juristisch-)universitären Domäne sei hierbei auf die Vorreiterrolle der Georg-August-Universität Göttingen mit ihrem Zertifikatsstudiengang „KI und Legal Tech (KILT)“ hingewiesen. Der Studiengang bereitet seine Absolventinnen und Absolventen auf die sich durch KI verändernde Lebensrealität vor, indem nicht nur die Kernkompetenzen im Umgang mit Künstlicher Intelligenz vermittelt, sondern diese auch anhand echter praktischer Tools erprobt werden. Die Lehrveranstaltungen werden hierbei von ausgewiesenen Expertinnen und Experten durchgeführt, die in ihrer beruflichen Praxis unmittelbar mit den aktuellen technischen und rechtlichen Entwicklungen befasst sind. Durch diesen Erwerb der AI Literacy wird den Studierenden eine echte Schlüsselqualifikation für ihr künftiges Berufsleben an die Hand gegeben. Der Zertifikatsstudiengang kann ohne Vorkenntnisse besucht werden und steht allen Studierenden der Georg-August-Universität offen.[181] Derartige Zertifikate sind ideal, um KI-Kompetenz nach Art. 4 KI-VO nachweisen zu können.[182] Die Kooperation mit Universitäten und Forschungseinrichtungen zur Vermittlung von Kenntnissen sowie gerade auch die Durchführung von Simulationen anhand von praktischen Anwendungsszenarien und Use Cases – insbesondere im disziplinär diversen Rahmen – werden in der einschlägigen Literatur mit Blick auf den Erwerb von KI-Kompetenz ausdrücklich hervorgehoben.[183]
Bezüglich der Kenntnisvermittlung kann ein mehrstufiger Ansatz verfolgt werden, indem zunächst eine einheitliche Schulung für den Personal-Stab angeboten wird, die je nach Bedarf durch weitere, spezialisierte Schulungen ergänzt wird.[184] Im Unternehmenskontext werden bereits derartige mehrstufige Schulungen

[179] Vgl. HK-KI-VO/Hilgendorf/Härtlein/*Hilgendorf*/*Härtlein*, 1. Aufl.2025, KI-VO Art. 4 Rn. 4ff.; siehe auch Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 24.
[180] Siehe etwa hierzu die Veröffentlichungen zum „Umgang mit KI-Modellen in Studium und Lehre“ seitens der *Georg-August-Universität Göttingen*, https://uni-goettingen.de/de/674738.html (Abruf am 20.04.2026, 19:52 Uhr); vgl. auch *Turbék*, Justiz und KI, Friedrich-Naumann-Stiftung für die Freiheit (Hrsg.), 2026, abrufbar unter https://shop.freiheit.org/#!/Publikation/2184 (Abruf am 09.10.2026, 15:40 Uhr), S. 32f.
[181] Georg-August-Universität Göttingen, Neuer Zertifikatsstudiengang: KI und Legal Tech, abrufbar unter https://uni-goettingen.de/de/neuer+zertifikatsstudiengang%3A+ki+und+legal+tech/698161.html (Abruf am 24.10.2026, 17:04 Uhr); *Reuß, Philipp M.*, Göttinger Forschungsstelle für e-Justiz, Automation und Digitalisierung (GOeJUSTICE), abrufbar unter https://reusz.eu/goejustice (Abruf am 05.08.2026, 21:30 Uhr).
[182] Martini/Wendehorst/*Wendehorst*, 2. Aufl. 2026, KI-VO Art. 4 Rn. 39.
[183] BeckOK KI-Recht/*Kaufmann*, 7. Ed. 25.7.2026, KI-VO Art. 4 Rn. 38.1.
[184] *Bunes*, in: Bernzen/Heinze/Steinrötter, DSRI Herbstakademie 2025, 1. Aufl. 2025, S. 376f.

empfohlen.[185] Eine derartige Stufung zeigt auch die Regelungstechnik der KI-VO auf. Art. 4 Abs. 1 S. 1 KI-VO erfasst die Nutzung und Entwicklung aller KI-Systeme i. S. v. Art. 3 Abs. 1 KI-VO und schreibt in der Folge aufgrund der Vielzahl der erfassten Fälle nur unkonkret vor, dass Maßnahmen ergriffen werden sollen, die Entwicklung von KI-Kompetenz zu unterstützen.[186] Art. 14 Abs. 1 KI-VO und Art. 26 Abs. 2 KI-VO adressieren den Bereich weitergehend (nur) für Hochrisiko-KI-Systeme. So müssen nach Art. 14 Abs. 1 KI-VO Hochrisiko-KI-Systeme so konzipiert und entwickelt werden, dass eine wirksame Aufsicht durch natürliche Personen erfolgen kann, was durch die Pflicht des Art. 26 Abs. 2 KI-VO komplementär adressiert wird, wonach die Betreiber verpflichtet sind, die Aufsicht nur entsprechend befähigten natürlichen Personen zu übertragen.[187] Vor diesem Hintergrund ist die voraussichtlich typischerweise erfolgende Vermittlung der Grundlagenkenntnisse aufgrund des Art. 4 Abs. 1 KI-VO (in der Regel) im Hochrisiko-Bereich um (die rollenspezifischen) Schulungen und Kompetenzen i. S. v. Art. 14 Abs. 1 i. V. m. Art. 26 Abs. 2KI-VO zu ergänzen.[188] (Auch) Im richterlichen Kontext bieten sich schriftlich fixierte, regelmäßig aktualisierte (unterstützende) Leitfäden oder Anleitungen an, die Richterinnen und Richtern niedrigschwellig zugänglich gemacht werden, etwa durch Verteilung mittels E-Mail sowie durch Abrufbarkeit im Intranet.[189] Ein derartiger Leitfaden sollte insbesondere typische systemseitige Eigenheiten, Risiken, Fehlerbilder und Prüfschritte vor der Übernahme systemseitiger Vorschläge beschreiben und zudem klarstellen, dass bei Auffälligkeiten der Systemausgaben der Abbruch der Arbeit mit dem informationstechnischen System angezeigt sein kann oder ist bzw. eine Pflicht zum eigenständigen richterlichen Eingreifen bestehen kann (vgl. in diesem Kontext auch Art. 14 Abs. 4 lit. d und e KI-VO).[190] Aufsichtsmaßnahmen in Bezug auf das KI-System können zudem zugleich adäquat an die Systemfunktion angepasst im System selbst eingebaut werden, um die Wahrnehmung der richterlichen Aufsichts- und

---

[185] Siehe etwa *Bitkom*, Umsetzungsleitfaden zur KI-Verordnung, Version 2.0, S. 154ff., abrufbar unter https://www.bitkom.org/Bitkom/Publikationen/Umsetzungsleitfaden-zur-KI-Verordnung (Abruf am 18.08.2026, 20:42 Uhr).

[186] *Hofer/Kirchmair*, ZfPC 2025, 270, 272.

[187] Vgl. *Hofer/Kirchmair*, ZfPC 2025, 270, 273f.

[188] *Hofer/Kirchmair*, ZfPC 2025, 270, 274.

[189] Vgl. *Bunes*, in: Bernzen/Heinze/Steinrötter, DSRI Herbstakademie 2025, 1. Aufl. 2025, S. 377; *Rollberg*, Algorithmen in der Justiz, 1. Aufl. 2020, S. 105ff.; *Schneider/Wundenberg/Sharaf*, ZHR 189 (2025), 259, 276; *Turbék*, Justiz und KI, Friedrich-Naumann-Stiftung für die Freiheit (Hrsg.), 2026, abrufbar unter https://shop.freiheit.org/#!/Publikation/2184 (Abruf am 09.10.2026, 15:40 Uhr), S. 30 empfiehlt für den Einsatz von KI in der Justiz institutionelle Leitlinien, die eine Vielzahl an Aspekten umfassen.

[190] Vgl. insoweit die im Rahmen des Konzepts zum „Umgang mit KI-Modellen in Studium und Lehre“ seitens der *Georg-August-Universität Göttingen* zur Verfügung gestellten Materialien, abrufbar unter https://uni-goettingen.de/de/674738.html (Abruf am 20.04.2026, 19:52 Uhr); Art. 14 Abs. 4 lit. d und e KI-VO verpflichtet Anbieter Hochrisiko-KI-Modelle den Betreibern derart zur Verfügung zu stellen, dass das Aufsichtspersonal adäquat in der Lage ist, das System nicht anzuwenden sowie in den Betrieb einzugreifen oder diesen derart zu unterbrechen, sodass es dem System ermöglicht ist, in einen sicheren Zustand zum Stillstand zu kommen, siehe hierzu BeckOK KI-Recht/*Buchner*, 7. Ed. 1.8.2026, KI-VO Art. 14 Rn. 58ff.; in diesem Zusammenhang sei auch auf die Betreiberpflicht zur Aussetzung der Verwendung des Systems nach Art. 26 Abs. 5 UAbs. 1 S. 2 KI-VO bei Grund zur Annahme eines Risikos nach Art. 79 Abs. 1 KI-VO [Risiko für Gesundheit oder Sicherheit oder Grundrechte von Personen] trotz betriebsanleitungsgemäßer Verwendung hingewiesen; vgl. auch Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 35.

Kontrollfunktion weiter zu unterstützen (vgl. auch die Anbieterpflicht in Art. 14 Abs. 3 KI-VO).[191]
Sinnvoll erscheint ferner die **Implementierung einer Pflicht, systemische Auffälligkeiten**, **unplausible Ergebnisse oder erkennbare Verzerrungen an eine zentrale Stelle zu melden**, damit Fehler nicht nur im Einzelfall korrigiert, sondern systematisch ausgewertet werden können.[192] Überwachungs- und Informationspflichten sind in verwandter Art und Weise auch in der KI-VO wesentliche Maßnahmen. Entsprechend des risikobasierten Ansatzes der KI-VO listet Art. 26 Abs. 5 UAbs. 1 KI-VO für den Hochrisiko-KI-Bereich verschieden intensive Maßnahmen auf, die betreiberseitig zu beachten sind.[193] Zunächst ist der Betrieb des Systems gemäß Art. 26 Abs. 5 UAbs. 1 S. 1 KI-VO anhand der Betriebsanleitung zu überwachen, zudem sind Anbieter ggf. gem. Art. 72 KI-VO (über die für sie wichtigen Angelegenheiten) zu informieren, um diesen die Beobachtung nach dem Inverkehrbringen zu erleichtern.[194] Haben Betreiber Grund zu der Annahme, dass die Verwendung gemäß der Betriebsanleitung dazu führen kann, dass das Hochrisiko-KI-System ein Risiko im Sinne des Art. 79 Abs. 1 KI-VO (ein solches ist ein Risiko für Gesundheit oder Sicherheit oder Grundrechte von Personen) birgt, so müssen sie gem. Art. 26 Abs. 5 UAbs. 1 S. 2 KI-VO unverzüglich den Anbieter oder Händler und die zuständige Marktüberwachungsbehörde i. S. v. Art. 3 Nr. 26 KI-VO – in Deutschland grundsätzlich die Bundesnetzagentur gem. § 2 Abs. 1 KI-MIG bzw. bei den vorliegend relevanten Hochrisiko-KI-Systemen die unabhängige KI-Marktüberwachungskammer bei der Bundesnetzagentur gem. § 2 Abs. 5 KI-MIG oder im Falle von KI-Systemen, die durch öffentliche Stellen der Länder in Verkehr gebracht, in Betrieb genommen oder verwendet werden nach § 2 Abs. 6 KI-MIG die nach Landesrecht zuständigen Behörden – informieren und die Verwendung dieses Systems aussetzen.[195] Im Fall eines „schwerwiegenden Vorfalls“ i. S. v. Art. 3 Nr. 49 KI-VO[196] sind gem. Art. 26 Abs. 5 UAbs. 1 S. 3 KI-VO zuerst der Anbieter und danach Einführer oder Händler sowie die zuständige Marktüberwachungsbehörde zu informieren. Im Falle der Nichterreichbarkeit des Anbieters hat der Betreiber gem. Art. 26 Abs. 5 UAbs. 1 S. 4 KI-VO die Pflichten des Anbieters selbst soweit möglich zu erfüllen.[197] Die Pflichten des Art. 26 Abs. 5 UAbs. 1 S. 1-4 KI-VO gelten gem. S. 5 nicht im Falle sensibler operativer Daten von Betreibern, die Strafverfolgungsbehörden i. S. v. Art. 3 Nr. 45 KI-VO sind, da die Offenlegung dieser Daten die Integrität des

[191] Siehe zu Art. 14 Abs. 3 KI-VO etwa BeckOK KI-Recht/*Buchner*, 7. Ed. 1.8.2026, KI-VO Art. 14 Rn. 40ff.
[192] Zur Vorteilhaftigkeit eines mehr-Ebenen-Systems mit einer selbstregulierenden Ausgestaltung siehe *Gaube/Langer/Miller/Baum/Dachselt/Feit/Gadiraju/Kaur/Keane/Landers/Laux/Liao/Lim/Onnasch/Schrills/Sonenberg/Tan/Tintarev/Xiao/Zhang*, Keeping an Eye on AI: A Framework for Effective Human Oversight of AI Systems, 2026, abrufbar unter https://doi.org/10.48550/arXiv.2605.16278 (Abruf am 24.10.2026, 23:00 Uhr), S. 7.
[193] BeckOK KI-Recht/*Denga*, 7. Ed. 1.8.2026, KI-VO Art. 26 Rn. 46.
[194] Martini/Wendehorst/*I. Eisenberger/Th. Binder*, 2. Aufl. 2026, KI-VO Art. 26 Rn. 35f.
[195] Bezüglich der Zuständigkeiten für Hochrisiko-KI-Systeme im vorliegend relevanten Bereich der richterlichen Entscheidungsfindung für die Marktüberwachungsbehörden sowie deren Tätigkeiten ist Art. 74 Abs. 8 KI-VO zu beachten.
[196] Ein solcher ist ein Vorfall oder eine Fehlfunktion bezüglich eines KI-Systems, das bzw. die direkt oder indirekt eine der nachstehenden Folgen hat: a) den Tod oder die schwere gesundheitliche Schädigung einer Person; b) eine schwere und unumkehrbare Störung der Verwaltung oder des Betriebs kritischer Infrastrukturen; c) die Verletzung von Pflichten aus den Unionsrechtsvorschriften zum Schutz der Grundrechte; d) schwere Sach- oder Umweltschäden.
[197] Martini/Wendehorst/*I. Eisenberger/Th. Binder*, 2. Aufl. 2026, KI-VO Art. 26 Rn. 38.

Strafverfahrens gefährden könnte.[198] Vor diesem Hintergrund erscheint die Implementierung eines breit angelegten **Informationsmanagementsystems** vorteilhaft, das eine zügige sowie effektive Erfassung, insbesondere von systemischen Auffälligkeiten, Unplausibilitäten, Verzerrungen sowie Risiken und Vorfällen, ermöglicht sowie deren Begegnung.[199] (Siehe zum Thema Monitoring, Fehlerkultur und kognitive Sicherungsmechanismen auch Abschnitt C.I.4.g.ee)
Auch sollten **Fortbildungsveranstaltungen, ggf. verpflichtend,** vorgesehen werden; im richterlichen Bereich bestehen solche derzeit bereits gem. § 22 Abs. 6 S. 2 GVG in Insolvenz- und Restrukturierungssachen sowie für Familienrichter gem. § 23b Abs. 3 S. 3–5 GVG.[200] Bei den Fortbildungen sollte auf die Best-Practices der Schulungsmaßnahmen aus der erst kürzlich abgeschlossenen E-Akten-Einführung zurückgegriffen werden. Auch bietet es sich an, die Fortbildungen aufzuzeichnen und niedrigschwellig abrufbar bereitzustellen sowie räumliche und zeitliche Synergien zu nutzen, d. h. etwa entsprechende Fortbildungsveranstaltungen als Zusatzmodule im Rahmen bereits terminierter Fortbildungsveranstaltungen zuzuschalten. Ebenso bieten sich FAQ-Seiten sowie Strukturen an, in denen problemlos auch triviale Fragen gestellt werden können. Eine anonyme Möglichkeit zur Fragestellung kann hierbei Hemmnissen entgegenwirken, derartige Fragen einzureichen und zu klären. Mit Blick auch auf die zügigen technologischen Fortschritte sollten die Schulungsinhalte hierbei regelmäßig aktualisiert werden.[201]
Ergänzend kommt eine **mehrstufige Qualitätskontrolle** in Betracht. Denkbar sind **stichprobenartige Nachprüfungen von Arbeitsergebnissen**, ein **Mehr-Augen-Prinzip** für besonders eingriffsintensive oder atypische Konstellationen (etwa in Anlehnung an § 348 Abs. 3 ZPO), **periodische Bias- und Fehlerratenanalysen**, (präventive) Prüfungen und wiederkehrende **Zertifizierungen** durch besonders sachkundige (institutionelle) Stellen sowie A/B-Tests mit verschiedenen Modellen bzw. Versionen und **Algorithmenkontrollen** einschließlich Protokollierung und Beweissicherung sowie die **Implementierung eines umfassenden Risikomanagementsystems** sowie Feedbacksystems.[202] **Kontinuierliche Routinevalidierungen der Systeme** sind ebenfalls vorteilhaft und können mithilfe von

---

[198] Siehe auch Art. 3 Nr. 38 KI-VO; BeckOK KI-Recht/*Denga*, 7. Ed. 1.8.2026, KI-VO Art. 26 Rn. 51; Martini/Wendehorst/*I. Eisenberger/Th. Binder*, 2. Aufl. 2026, KI-VO Art. 26 Rn. 39.
[199] Vgl. auch Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 126.
[200] Vgl. *Reformkommission Zivilprozess der Zukunft*, Abschlussbericht der Reformkommission Zivilprozess der Zukunft im Auftrag des 3. Digitalgipfels der Justizministerinnen und Justizminister des Bundes und der Länder, abrufbar unter https://www.bmjv.de/SharedDocs/Publikationen/DE/Fachpublikationen/2025_Abschlussbereicht_Reformkomission_Zivilprozess.html (Abruf am 20.04.2026, 21:09 Uhr), S. 236f.
[201] Vgl. *Hofer/Kirchmair*, ZfPC 2025, 270, 273.
[202] *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 368f., 373f.; vgl. *Rollberg*, Algorithmen in der Justiz, 1. Aufl. 2020, S. 108ff.; *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 292ff.; mit Blick auf den behördlichen Kontext vgl. *Martini/Nink*, NVwZ-Extra 10/2017, 1, 12f.; Bomhard/Pieper/Wende/*Bomhard/Wietzke*, 1. Aufl. 2025, KI-VO Art. 14 Rn. 21f.; siehe auch Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 125f.

Massendatenauswertungen statistisch relevante Zusammenhänge in der Verarbeitung zutage fördern und damit möglicherweise diskriminierende Filterkriterien des informationstechnischen Systems aufdecken; in der Folge können entsprechend ausgebildete Experten die informationstechnischen Systeme mittels Kontrollalgorithmen prüfen sowie mittels Testverfahren auf diskriminierende Strukturen hin untersuchen, sie ggf. identifizieren und korrigieren.[203] Dabei sollte nicht nur auf offen diskriminierende Ergebnisse geachtet werden; zu achten ist auch auf weitere Verzerrungen, etwa infolge unausgewogener Datenauswahl, Systemmodellierung – wie etwa auch auf die Entscheidungsbaummodellierung.[204] Dies gilt besonders bei Vorliegen von Rechts- oder Rechtsprechungsänderungen, die in den sachlich-inhaltlichen Einsatzszenarien von Systemen nachvollzogen werden müssen.
Des Weiteren ist die **Benutzeroberflächengestaltung** des informationstechnischen Systems zu beachten, insbesondere mit Blick auf den Automation Bias, bei dem u. a. die konkrete Hinweisplatzierung sowie die Bereitstellung von Informationen im Gegensatz zu bloßen Empfehlungen Gegenmaßnahmen darstellen können.[205]
Sicherungsmaßnahmen gegen eine strukturelle Besorgnis der Befangenheit infolge **Bias** – wobei Voraussetzung sein dürfte, dass diese potentiell Auswirkungen auf das konkrete Ergebnis zeitigen – können zudem systemseitig vorgesehene Regeln sein, die die Unvoreingenommenheit sicherstellen, nachträgliche Kontrollen sowie Ablehnungsmöglichkeiten analog der Besorgnis der Befangenheit.[206] Im Besonderen mit Blick auf die Ablehnung aufgrund einer Besorgnis der Befangenheit könnte für den entscheidungs*ersetzenden* Einsatz eine gesetzliche Klarstellung in den §§ 41 ff. ZPO erwogen werden, um der Notwendigkeit der Analogie zuvorzukommen und Rechtssicherheit für derartige Fälle zu schaffen. Mit Blick auf die Kontrolle der Bias-Freiheit bietet sich wiederum die Einbindung beispielsweise in Form einer zyklischen Überprüfung der Ausgaben durch entsprechend sachkundige Stellen an.[207] Im Rahmen des entscheidungs*unterstützenden* Einsatzes informationstechnischer Systeme behält der menschliche Richter bzw. die menschliche Richterin die Letztentscheidung, das System wird hier nur unterstützend tätig. Insoweit ist der derzeitige Bezugspunkt der Regelungen über die Ausschließung und Ablehnung der Gerichtspersonen – als menschliche Personen – unverändert. Es besteht daher auch kein grundlegender Handlungsbedarf, diese Regelungen auf den Fall des entscheidungs*unterstützenden* Einsatzes informationstechnischer Systeme anzupassen. Mit Blick auf die Bias-Thematik könnte zwar eine gesetzliche Ergänzung bzw. Klarstellung – z. B. durch Ergänzung des Ausschlussgründe-Katalogs kraft Gesetzes nach § 41 ZPO – in Betracht gezogen werden, gleichwohl dürften entsprechende Fälle regelmäßig bereits durch den bestehenden Normbestand abgedeckt sein. So können Unsachlichkeiten und eine mangelhafte Art und Weise der Verfahrensführung bereits derzeit den Ablehnungsgrund der Besorgnis der Befangenheit begründen.[208] Hierunter könnten mit Blick auf den Einsatz entscheidungs*unterstützender* informationstechnischer Systeme beispielsweise

[203] *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 369; vgl. *Martini/Nink*, NVwZ-Extra 10/2017, 1, 12f., mit Blick auf den behördlichen Kontext.
[204] *Kernchen*, Automatisierung und richterliche Rechtsanwendung, 1. Aufl. 2025, S. 137ff.
[205] *Goddard/Roudsari/Wyatt*, Automation bias: a systematic review of frequency, effect mediators, and mitigators, Journal of the American Medical Informatics Association 19 (2012), No. 1, S. 121, s. a. S. 125 m. w. N.; vgl. auch *Dötterl*, RDi 2026, 414, 419.
[206] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 356ff. m. w. N.
[207] *Martini/Nink*, NVwZ-Extra 10/2017, 1, 12.
[208] Musielak/Voit/*Heinrich*, 23. Aufl. 2026, ZPO § 42 Rn. 11, 13.

Sachverhalte fallen, in denen ein Richter bzw. eine Richterin auf den parteiseitigen Hinweis in der mündlichen Verhandlung, dass ein Vortrag bereits im schriftlichen Verfahren vorgetragen worden sei, entgegnet, dass dies ja wohl offensichtlich nicht stimmen könne, da das informationstechnische Unterstützungssystem dies nicht anzeige (ggf. um die Aussage ergänzt, dass man nicht derart klar vor Gericht lügen solle). Im Zusammenhang mit dem entscheidungs*ersetzenden* Einsatz informationstechnischer Systeme könnte eine entsprechende Anwendung für das System vergleichbar § 49 ZPO für die Urkundsbeamten angeordnet werden – sei es, da man das System nicht unter den Richterbegriff im Rahmen der §§ 41 ff. ZPO subsumieren wollte oder schlicht aus Klarstellungsgründen. Zur Vermeidung des „bösen Scheins“[209] der Parteilichkeit könnte zudem klargestellt werden, dass bei einer Bereitstellung eines Systems durch einen (privaten) Dritten, in Fällen, in denen dieser eine am Rechtsstreit beteiligte Partei ist, das informationstechnische System von der Ausübung des Richteramts ausgeschlossen ist. In Fällen begründeter Ablehnung (vgl. § 42 Abs. 1 ZPO) sowie eines Ausschlusses des informationstechnischen entscheidungs*ersetzenden* Systems von der Ausübung des Richteramts sollte eine Aussteuerung an einen menschlichen Richter bzw. eine menschliche Richterin vorgesehen werden. Ein weiteres für die Entscheidung des konkreten Falls taugliches System wird präsumtiv nicht vorliegen und wäre auch aus Gründen der Kosten von Parallelentwicklungen nicht vorzugswürdig, sodass die Option der Abgabe des Falles an ein anderes informationstechnisches System abzulehnen ist. Auch aus Akzeptanzgesichtspunkten erscheint es vorteilhaft, die Parteien nicht erneut einem informationstechnischen System zu unterwerfen, nachdem ein anderes bereits ausgeschlossen oder abgelehnt wurde, sondern dem Normalfall des menschlichen Richters bzw. der menschlichen Richterin.
Es sollten ferner **dienstrechtliche Maßnahmen** vorgesehen werden für Fälle, in denen Richterinnen und Richter (offensichtlich und jeden Zweifels erhaben) uneingeschränkt sowie unreflektiert Ergebnisse eines informationstechnischen Systems im Rahmen ihrer Entscheidungsfindung übernehmen bzw. in entsprechender Weise gegen ihre Aufsichts- und Kontrollpflicht verstoßen.[210] Soweit für gravierende Verstöße Maßnahmen dienstrechtlicher Natur vorgesehen sind, sollten deren Voraussetzungen sowie die Art der Sanktionen hinreichend genau bestimmt sein und kommuniziert werden, damit ein klarer und für den Einzelnen rechtssicherer Rechtskorridor im Rahmen des Einsatzes informationstechnischer Systeme bei der richterlichen Entscheidungsfindung besteht und kein abschreckender Effekt eintritt, der den Einsatz dieser informationstechnischen Systeme faktisch unterbindet. Bei der Ausdifferenzierung von derartigen Maßnahmen sollten insbesondere derartige Auswirkungen im Auge behalten werden sowie die rechtspolitische Fragestellung, inwieweit der Einsatz derartiger Systeme dem Grunde nach gewünscht ist.

[209] Vgl. MüKoZPO/*Stackmann*, 7. Aufl. 2025, ZPO § 42 Rn. 7.
[210] Zur Zulässigkeit solcher Maßnahmen in diesen Fällen vgl. *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 193; *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 300; *Arbeitsgruppe „Einsatz von KI und algorithmischen Systemen in der Justiz“*, Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier zur 74. Jahrestagung der Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs vom 23. bis 25. Mai 2022 in Rostock, abrufbar unter https://www.justiz.bayern.de/media/images/behoerden-und-gerichte/oberlandesgerichte/nuernberg/einsatz_von_ki_und_algorithmischen_systemen_in_der_justiz.pdf (Abruf am 15.04.2026, 18:58 Uhr), S. 13.

(4) Ausschluss inhaltlicher Kontrolle richterlicher Dokumente durch Verwaltung oder Dritte

Bezüglich der zwingend auszuschließenden inhaltlichen Kontrolle richterlicher Dokumente durch Verwaltung oder Dritte empfiehlt sich eine **weitestgehende Trennung technischer und organisatorischer Art** der externen Zugriffsrechte sowie richterlichen Arbeitssphäre;[211] maßgeblich sind in diesem Zusammenhang differenzierte Protokollierungs- sowie Zugriffsrechtekonzepte, klare und transparente Regeln für die Zugriffsvoraussetzungen und -sicherheitsvorkehrungen, das zugreifende Personal und zu welchen Zwecken auf Inhalts- sowie Protokolldaten zugegriffen werden darf sowie diese ausgewertet und verwendet werden dürfen.[212] Die Kontrollmöglichkeit sowie Kontrolle bezüglich der Beachtung dieser Regeln muss gegeben sein; die Verwertung der Inhalts- sowie Protokolldaten für dienstaufsichtsrechtliche Zwecke ist grundsätzlich auszuschließen und etwaige Ausnahmen klar zu fassen.[213] Die Schutzvorkehrungen müssen bei Fremddatenspeicherungen derart wirksam sein, dass schon der „böse Schein“ eines unbefugten Zugriffs auf gerichtliche bzw. richterliche Dokumente ausgeschlossen wird.[214]

(5) Diskriminierungssensible Auswahl und Kontrolle der Trainingsdaten

Mit Blick auf die Auswahl und Gestaltung der Trainingsdaten sollten Qualitätskontrollen während der Systementwicklung erfolgen, damit unsachgemäße menschliche Einflüsse mit Ergebnisrelevanz nicht in systemseitige Resultate einfließen.[215] Informationstechnische Systeme maschinellen Lernens (auf Grundlage der Methode überwachten Lernens[216]) können – unter der Einhaltung bestimmter Voraussetzungen – als tauglich angesehen werden, die Gesetzesbindung zu wahren und die Rechtsanwendung abzubilden.[217] Damit ist der Einsatz bei informationstechnischen Systemen bei der richterlichen Entscheidungsfindung möglich. Hierzu bedarf es geeigneter, repräsentativer sowie entsprechend aufbereiteter Daten(sätze).[218] Datenauswahl und Modellierung des Systems sind wesentliche Faktoren, die zu Verzerrungen bzw. Bias führen können und sind daher als wesentliche Faktoren zu betrachten; die Verzerrungen sind oftmals durch eine unterkomplex verfügbare bzw. ausgewählte Stichprobe sowie Merkmale bedingt.[219] Eine einheitliche Lösung gegen systemseitige Diskriminierungen und Bias gibt es nicht, gleichwohl werden technische Einwirkungen und Systemgestaltungen als Möglichkeiten erachtet, Linderung zu verschaffen.[220] Beispielsweise können Programmierer die Anknüpfung an bestimmte

[211] Vgl. zum Kontext der sog. „Netzklage(n)“ *Süptitz/Eymann*, IT-Outsourcing für die Justiz: Ein Eingriff in die richterliche Unabhängigkeit?, in: Räckers/Halsbenning/Rätz/Richter/Schweighofer (Hrsg.), Digitalisierung von Staat und Verwaltung, Lecture Notes in Informatics (LNI), Bonn 2019, S. 15, 21ff.
[212] *Berlit*, Richterliche Unabhängigkeit und elektronische Akte, JurPC Web-Dok. 77/2012, Abs. 39f.
[213] *Berlit*, Richterliche Unabhängigkeit und elektronische Akte, JurPC Web-Dok. 77/2012, Abs. 40.
[214] *Berlit*, Richterliche Unabhängigkeit und elektronische Akte, JurPC Web-Dok. 77/2012, Abs. 45.
[215] *Kernchen*, Automatisierung und richterliche Rechtsanwendung, 1. Aufl. 2025, S. 137, 140.
[216] Vgl. *Quack*, Automation zivilrechtlicher Entscheidungsfindung, 1. Aufl. 2025, S. 122ff., 146.
[217] Siehe hierzu eingehend und auch mit Hinweis auf andere Ansichten *Kernchen*, Automatisierung und richterliche Rechtsanwendung, 1. Aufl. 2025, S. 122ff.
[218] Vgl. *Kernchen*, Automatisierung und richterliche Rechtsanwendung, 1. Aufl. 2025, insbesondere S. 127ff., 155ff.
[219] *Kernchen*, Automatisierung und richterliche Rechtsanwendung, 1. Aufl. 2025, S. 137; *Wischmeyer*, AöR 143 (2018), 1, 27.
[220] *Wolff*, Algorithmen als Richter, 1. Aufl. 2022, S. 78.

Merkmale, wie Geschlecht und „Rasse“, oder Datenerhebungen in gewissem Umfang ausschließen.[221] Als weitere Verfahren zur Bias-Bekämpfung sind statistische Analysen zu nennen, beispielsweise des „Disparity Impacts“, die potenziell diskriminierende Verzerrungen aufdecken können, Ursache-Wirkungs-Analysen, bei denen beispielsweise mittels minimalinvasiver Änderung von Datenpunkten und Monitoring der Veränderungen der Systemergebnisse, verzerrungsrelevante Datenpunkte identifiziert werden können.[222] Zudem kommen inhaltliche Analysen von Daten sowie Analysen mittels Lexika oder Stimmungsanalysen in Betracht, behaftete Begriffe oder Grundstimmungen zu identifizieren, und die Hinzuziehung spezialisierter Anbieter (z. B. AI Fairness 360 [IBM] oder Fairness Indicators).[223] Mit Blick auf die Trainingsdatenauswahl und -gestaltung erscheint ein dokumentiertes Datenmanagement vorteilhaft, das eine Transparenz über die Intention, Zusammensetzung, den Erhebungsprozess, Datenkuratierung sowie Verwendung und Pflege der Daten offenlegt.[224] Freilich sind die Trainingsdaten insbesondere darauf zu untersuchen, ob sie den Datensatz mit historischen Bias (etwa Stereotype) oder repräsentationsbezogenen Bias (Unterrepräsentation der relevanten Bevölkerungsgruppe) infizieren.[225] Es ist (daher) Wert auf eine ausreichend breite Datenbasis, die Bereinigung des Datenbestandes, die Repräsentativität, Fehlerfreiheit und Aktualität des Datensatzes zu legen.[226] Für juristische Daten sollten bei der Annotation möglichst klare Annotationsrichtlinien eingesetzt werden, um sowohl die Konsistenz der Annotation über die verschiedenen Bearbeiter hinweg als auch die Rechtskonformität – insbesondere hier mit Blick auf die Diskriminierungsfreiheit – zu gewährleisten. Jedoch sind auch nicht nur primär in Trainingsdaten begründete Diskriminierungs-Gefahren zu betrachten. Sog. statistische Diskriminierungen können aufgrund der Funktion informationstechnischer Systeme maschinellen Lernens auftreten, indem personenbezogene Merkmale als Stellvertretermerkmale, die in einem Korrelationsverhältnis zur relevanten Eigenschaft stehen, anstatt der relevanten

---

[221] *Wischmeyer*, AöR 143 (2018), 1, 27.
[222] Bomhard/Pieper/Wende/*Bomhard/Wietzke*, 1. Aufl. 2025, KI-VO Art. 10 Rn. 48.
[223] Bomhard/Pieper/Wende/*Bomhard/Wietzke*, 1. Aufl. 2025, KI-VO Art. 10 Rn. 48.
[224] *Gebru/Morgenstern/Vecchione/Vaughan/Wallach/Daumé III/Crawford*, Datasheets for Datasets, abrufbar unter https://doi.org/10.48550/arXiv.1803.09010 (Abruf am 23.09.2026, 20:42 Uhr), S. 1ff.; vgl. auch Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 4; siehe auch *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 27f.
[225] *Suresh/Guttag*, A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle, Proceedings of the 1st ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '21) 2021, S. 4.
[226] Vgl. *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 393; siehe zu Aspekten des Art. 10 KI-VO etwa Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 26ff., 47ff.; siehe zudem auch a. a. O. S. 126; siehe auch *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 29f.

Eigenschaft den Ausschlag zum (falschen) Ergebnis bringen.[227] Auch gegen diese mittelbaren Diskriminierungen, als mittelbare Folge der Designentscheidung der Programmierer und Systementwickler, sind Kontrollmechanismen vorzusehen.[228] Freilich sind auch Bias in der Entwicklung sowie in der Anwendung der KI-Systeme umfassend zu beachten und entsprechend zu adressieren.[229] Letztlich sollte die Bias-Problematik auf allen Ebenen der Technik-Architektur adressiert werden.[230] Hierbei sollte jedoch auch beachtet werden, dass Ungleichbehandlungen zwar streng zu adressieren sind, gleichwohl solche auch mit hinreichenden Umständen gerechtfertigt werden könnten.[231] Auch sollte beachtet (und die Aspekte miteinander in einen gebührenden Ausgleich gebracht) werden, dass verschiedene rechtliche Gewährleistungen einander beeinflussen können, insbesondere Datenschutz und Antidiskriminierung; der Verzicht auf bestimmte Datenarten kann unter dem Aspekt der Antidiskriminierung durchaus vorteilhaft sein, jedoch kann ein Mangel von spezifizierten Daten auch dazu führen, dass Systeme maschinellen Lernens Entscheidungen primär anhand von Stellvertretermerkmalen treffen, die ihrerseits wieder zu Verzerrungen und (diskriminierenden) Beeinträchtigungen führen, die gerade vermieden werden sollten.[232]
Zusätzlich kann eine **fortlaufende Daten- und Ergebnisevaluation** (mit Stichprobenpflichten zur frühzeitigen Erkennung und Ausbesserung von Fehlern) vorgesehen werden.[233] Auch sind hierzu interdisziplinär-wissenschaftlich begleitete Testläufe zu nennen sowie spezielle, ggf. gesonderte, Prüfprogramme für Randfallkonstellationen, da diese der Gefahr ausgesetzt sind, dass die Aussagekraft der vorhandenen Daten infolge der geringen Trainingsdatenlage systemseitig überschätzt wird.[234] Im Sinne einer stetigen Evaluation und Optimierung[235] sollte bei richterlichen Korrekturen zumindest die Möglichkeit bestehen, diese an die Datenpflege zurückzuspielen, damit Fehler systematisch erfasst und ausgebessert werden können.

cc. Zusammenfassung

Die menschliche Aufsicht und Kontrolle ist eine der zentralen roten Linien für jeden Einsatz informationstechnischer Systeme bei der richterlichen Entscheidungsfindung. Demnach sind Richterinnen und Richter umfassend in Beschaffung, Betrieb und Überwachung des Systems einzubinden. Beim entscheidungsersetzenden Einsatz und bei der Berührung des Kernbereichs der Rechtsprechung im entscheidungsunterstützenden Einsatzszeanrio sind die Richterinnen und Richter zudem in die Entwicklung einzubinden. Die Beauftragung Privater oder der Verwaltung mit der Entwicklung der Systeme ist möglich, bei Wahrung hinreichenden richterlichen Einflusses auf Entwicklung, Training, Konfiguration und Betrieb sowie Überwachung und Kontrolle der Funktionsfähigkeit. Eine inhaltliche Kontrolle richterlicher Dokumente durch Dritte ist auszuschließen. Ein Einsatz scheidet aus, wenn ein ergebnisrelevanter

[227] *Kernchen*, Automatisierung und richterliche Rechtsanwendung, 1. Aufl. 2025, S. 139f.
[228] *Kernchen*, Automatisierung und richterliche Rechtsanwendung, 1. Aufl. 2025, S. 139ff.
[229] Vgl. ausführlich Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 10 Rn. 37ff., 48ff.
[230] *Wischmeyer*, AöR 143 (2018), 1, 28f. mit weiteren Einzelbeispielen.
[231] *Wischmeyer*, AöR 143 (2018), 1, 30.
[232] *Wischmeyer*, AöR 143 (2018), 1, 31f.
[233] *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 373f.
[234] Hinsichtlich des Gefahren-Aspekts *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 373, 393.
[235] *Kernchen*, Automatisierung und richterliche Rechtsanwendung, 1. Aufl. 2025, S. 225.

Bias nicht ausgeschlossen werden kann bzw. die menschliche Kontrolle nicht gewährleistet ist.
Die Etablierung eines **richterlichen Governance-Modells**, in dem die richterliche Aufsicht und Kontrolle institutionell abgesichert ist, ist unbedingt anzuraten. Für die Ausgestaltung bestehen mehrere Möglichkeiten, wie z. B. interdisziplinär besetzte „Automatisierungskammern“, Automatisierungsinstitutionen oder Verbundarbeiten mit Externen. Ein einziges Modell, das für alle Einsatzszenarien angemessen ist, besteht nicht, vielmehr ist das passende Governance-Modell anhand des konkret geplanten Einsatzes adäquat zu entwickeln und zu etablieren, mittels der Faktoren des funktionalen Charakters der Tätigkeit, des Automatisierungsgrads und der Komplexität des Anwendungsfeldes. Mit der Zunahme der Faktoren dieser dreigliedrigen Funktionsmatrix steigen die Anforderungen an die richterliche Aufsicht und Kontrolle.
Hinsichtlich der Entwicklung und des Betriebs des informationstechnischen Systems kommt einerseits eine **Ansiedelung bei der Exekutive bzw. bei privaten Dritten** in Betracht. In diesem Falle bedarf es insbesondere klarer Regelungen der Zugriffsrechte sowie effektiver Kontroll- und Mitsprachemöglichkeiten für die Judikative. Vorzugswürdig erscheint hierbei u. a. ein (institutionell ausgeformter) richterlicher IT-Beauftragter. Ist eine hinreichende Aufsicht und Kontrolle nicht sicherzustellen, scheidet die externe Ansiedelung aus. Auf der anderen Seite kommt eine originäre Ansiedelung der Entwicklung und des Betriebs der informationstechnischen Systeme bei der Judikative selbst in Betracht. Diese kann ggf. auch nach Einsatzszenario abgestuft verfolgt werden. Mit steigendem Funktionsgrad wächst die Vorzugswürdigkeit einer vollständigen **Ansiedelung bei der Judikative**. Hinsichtlich der Thematik der Ansiedelung des informationstechnischen Systems ist vor diesem Hintergrund zu empfehlen, die Entwicklung und den Betrieb umso eher bei der Judikative anzusiedeln, je stärker das informationstechnische System inhaltsbezogene Aufgaben wahrnimmt. Dies ist auch mit Blick auf das Prinzip der technologischen Souveränität vorteilhaft.
Für die **Überwachung der Arbeitsergebnisse** und den **Umgang mit Bias** sowie sonstigen Fehlmustern ist insbesondere die Vermittlung der nötigen rechtlichen und technischen Kenntnisse wesentlich, durch bedarfsgerechte Schulungen und Fortbildungen. Dies spiegelt sich auch in den Anforderungen der KI-VO wider. Sinnvoll erscheinen u. a. ein frühzeitiger Schulungsansatz, mehrstufige Ausgestaltungen der Schulungen, die Implementierung von Leitfäden sowie ggf. verpflichtende Fortbildungsveranstaltungen. Zudem erscheint die Implementierung einer Pflicht, systemische Auffälligkeiten, unplausible Ergebnisse oder erkennbare Verzerrungen an eine zentrale Stelle zu melden, eingebettet in ein breit angelegtes Informationsmanagementsystem sinnvoll. Ergänzend kommt eine mehrstufige Qualitätskontrolle in Betracht; auch die **Benutzeroberflächengestaltung** sollte mit Blick insbesondere auf den Automation Bias Augenmerk erfahren sowie dienstrechtliche Maßnahmen bei offenkundigen Verletzungen der Aufsichts- und Kontrollpflicht. Für den entscheidungsersetzenden Einsatz könnte eine gesetzliche Klarstellung in den §§ 41 ff. ZPO erwogen werden; im entscheidungsunterstützenden Einsatz behält der menschliche Richter bzw. die menschliche Richterin die Letztverantwortung, hier besteht insoweit kein grundlegender Handlungsbedarf.
Für den Ausschluss einer inhaltlichen Kontrolle richterlicher Dokumente durch die Verwaltung oder Dritte empfiehlt sich eine weitestgehende Trennung technischer und organisatorischer Art der externen Zugriffsrechte sowie richterlichen Arbeitssphäre. Bei Fremddatenspeicherungen müssen Schutzvorkehrungen derart wirksam sein, dass schon der „böse Schein“ eines unbefugten Zugriffs auf gerichtliche bzw. richterliche Dokumente ausgeschlossen wird.

Die Trainingsdaten für Systeme sollten bereits während der Entwicklung hinsichtlich der Qualität kontrolliert werden. Wert zu legen ist auf einen hinreichend geeigneten Datensatz. Eine einheitliche Lösung gegen systemseitige Diskriminierungen und Bias existiert nicht, die Problematik sollte auf allen Ebenen der Technik-Architektur adressiert werden. Die Trainingsdaten sind diskriminierungssensibel auszuwählen. Ein dokumentiertes Datenmanagement und fortlaufende Evaluationen sind hierbei vorteilhaft.

b. Rote Linien für einen entscheidungs*ersetzenden* Einsatz sowie Handlungsleitlinien

Neben den übergreifenden roten Linien, die unter C.II.2.a bereits eingehend betrachtet worden sind, ergeben sich für den spezifischen Einsatz als entscheidungs*ersetzendes* System weitere rote Linien sowie Handlungsoptionen und -empfehlungen.

aa. Rote Linien entscheidungs*ersetzende*r Einsatz

Aufgrund der gegenwärtigen technischen Limitierungen der Leistungsfähigkeit informationstechnischer Systeme können diese rechtlich zulässig derzeit nur in einem beschränkten Rahmen eingesetzt werden.[236] Ein solcher Einsatz ist nur in **hinreichend einfach gelagerten Fällen** zulässig, in denen sich überschaubare Tatsachen- und Rechtsfragen von begrenzter Komplexität stellen sowie nur geringe Auslegungsspielräume bestehen und deswegen der Weg, auf dem das informationstechnische System zu seiner Entscheidung gelangt, einer mathematisch-logischen Rechenoperation gleichkommt.[237] Ein solch einfach gelagerter Fall scheidet insbesondere aus, wenn das informationstechnische System komplexe Abwägungsentscheidungen treffen müsste.[238]
Die vorstehend genannte rote Linie ergibt sich aus zahlreichen Verbürgungen und insbesondere aus der richterlichen Unabhängigkeit[239], der Menschenwürdegarantie gemäß Art. 1 Abs. 1 GG[240] und Art. 1 GRCh[241], dem Gleichbehandlungsgrundsatz gemäß Art. 3 GG[242], Art. 14 EMRK[243] und Art. 20–26 GRCh[244], dem Justizgewährungsanspruch gemäß Art. 2 Abs. 1, 20 Abs. 3 GG[245] und dem Rechtsstaatsprinzip[246] sowie dem Anspruch auf ein faires Verfahren gemäß Art. 2 Abs. 1, 3 Abs. 1, 20 Abs. 3 GG[247] und aus Art. 6 Abs. 1 EMRK[248] sowie Art. 47 Abs. 1, Abs. 2 GRCh[249].

---

[236] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 87ff., 106ff., 115ff., 143ff., 170ff., 229ff., 240ff., 255f., (258ff.,) 265ff., 334f., 338f., 344, 356ff.
[237] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 87ff., 106ff., 115ff., 143ff., 170ff., 229ff., 240ff., 255f., (258ff.,) 265ff., 334f., 338f., 344, 349ff., 356ff.
[238] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 115ff., 170ff., 229ff., 242ff., 338f., 356ff.
[239] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 170ff.
[240] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 87ff.
[241] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 265ff.
[242] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 106ff.
[243] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 255f.
[244] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 265ff.
[245] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 115ff.
[246] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 143ff.
[247] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 355ff.
[248] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 242ff.
[249] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 265ff.

Liegt ein hinreichend einfach gelagerter Fall nicht vor, etwa weil die Fallentscheidung eine komplexe Abwägungsentscheidung erfordert, darf aufgrund der vorstehenden Verbürgungen ein entscheidungs*ersetzender* Einsatz nicht erfolgen. Es muss daher jederzeit eine **systemseitige Aussteuerung an den menschlichen Richter (auch auf Antrag)** möglich sein, wenn sich herausstellt, dass kein einfach gelagerter Fall in diesem Sinne vorliegt.[250]
Darüber hinaus müssen systemseitige **Halluzinationen**[251] nach Möglichkeit vermieden werden. Wo dies nicht hinreichend sichergestellt ist, kann ebenfalls ein entscheidungs*ersetzender* Einsatz nicht erfolgen.[252] Halluzinationen geraten insbesondere mit der Menschenwürdegarantie gem. Art. 1 GG[253], dem Gleichbehandlungsgrundsatz gemäß Art. 3 GG[254], Art. 14 EMRK[255] und Art. 20–26 GRCh[256] in Konflikt, da es dadurch zu Ungleichbehandlungen kommen kann. Zudem beeinträchtigen sie das Rechtsstaatsprinzip gemäß Art. 20 Abs. 3 GG[257] und die Art. 92, 101 Abs. 1 Satz 2GG[258] sowie Art. 6 Abs. 1 EMRK[259].

bb. Handlungsoptionen und -empfehlungen

Die Einhaltung der rechtlichen Einsatzgrenzen wird aufgrund der Limitierungen der gegenwärtigen Systeme **derzeit nur von regelbasierten Systemen in hinreichend einfach gelagerten Fällen gewährleistet**.[260] In derartigen Fällen – insbesondere Massenverfahren wie beispielsweise bei einigen Fluggastrechteverfahren – können bereits jetzt verfügbare Systeme einfach erfassbare Daten fehlerfrei und schneller als der Mensch verarbeiten[261] und zu erheblichen Rationalisierungsgewinnen führen. Grundlegend – insbesondere mit Blick auf den entscheidungsersetzenden Einsatz, aber auch mit Blick auf den entscheidungsunterstützenden Einsatz informationstechnischer Systeme – sollte beachtet werden, dass zwar die Nutzung

[250] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 143ff., 229ff., 249, 291ff., 334f., siehe auch S. 90, 344 sowie im Falle einer Ausgestaltung des entscheidungsersetzenden Systems ähnlich dem Mahnverfahren u. a. S. 117, 231f.
[251] Siehe zu der Begrifflichkeit insbesondere *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 29, 89 m. w. N.
[252] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 87ff., 106ff., 143ff., 170f., 189f., 212ff., 241ff., 255f., 265ff., 333ff.
[253] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 87ff., 106ff.
[254] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 106ff.
[255] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 255f.
[256] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 265ff.
[257] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 143ff.
[258] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 170f., 212ff.
[259] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 241ff.
[260] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 89f.; einen ersetzenden Einsatz informationstechnischer Systeme für richterliche Rechtsprechungstätigkeit ablehnend vgl. statt vieler nur Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 9f., siehe aber a. a. O. S. 10f., wo ein Szenario im Rahmen eines entscheidungsunterstützenden Einsatzes regelbasierter Systeme als sinnvoll befürwortet wird, das materiell den vorliegenden Anforderungen bzw. der vorliegenden Ausgestaltung der Möglichkeit des Einsatzes entscheidungsersetzender informationstechnischer Systeme im Anwendungsfall entspricht.
[261] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 88, 107, 171ff.

informationstechnischer Systeme Risiken birgt, die Nichtnutzung informationstechnischer Systeme bei der richterlichen Entscheidungsfindung jedoch nicht ohne Kosten ihrerseits einhergeht; eine Nutzung hat insbesondere Potential, Zeit sowie Kosten zu ersparen und insbesondere Personalressourcen anderweitig verfügbar zu machen, ihre Nichtnutzung geht insbesondere mit den Kosten der Nichtergreifung dieser entlastenden Aspekte einher.[262]

(1) Identifikation hinreichend einfach gelagerter Fälle und (Teil-)Einsatzszenarien

Eine **erste und auch notwendige Handlungsoption** besteht darin, die **relevanten und möglichen Einsatzbereiche zu identifizieren**, in denen ein entscheidungsersetzender Einsatz in Betracht kommt. Insbesondere bei Massenverfahren verspricht der Einsatz informationstechnischer Systeme besondere Effizienzgewinne, aber auch bei Verfahren im Bereich des Verbraucherrechtedurchsetzungsgesetzes (§§ 14, 20 VDuG), bei der Kostenfestsetzung nach einer Kostengrundentscheidung gem. der §§ 103, 104 ZPO sowie bei der Abwicklung von Verkehrsunfällen bestehen Potentiale.[263]
Es ist hierbei **zu empfehlen**, im Rahmen interdisziplinär besetzter – und Richterinnen und Richter notwendigerweise mit einschließender – Arbeitsgruppen (zu Entwicklungs-, Beschaffungs-, Betriebs- und Überwachungsmöglichkeiten bereits Abschnitt C.II.2.a.bb(1)) entsprechende Einsatzfelder zu ermitteln und die möglichen Entscheidungskriterien und Sachverhaltskonstellationen in einem möglichen Entscheidungsbaum präzise herauszuarbeiten.
Dabei kann es auch durchaus förderlich sein, bei einem entsprechend zu erwartenden Effizienzgewinn, **informationstechnische Systeme bereits für Teilbereiche** zu entwickeln,[264] beispielsweise nur für bestimmte, automatisiert abzuhandelnde einfache Fallkonstellationen einer Verfahrenskategorie und die Beibehaltung einer menschlichen Entscheidung in komplexeren Sachverhalten dieser Kategorie. So lassen sich beispielsweise Ansprüche aus der Fluggastrechte-VO wegen Nichtbeförderung (Art. 4 Fluggastrechte-VO) durchaus gut automatisiert behandeln und könnten einem automatisierten Entscheidungssystem nach regelbasierter Funktionalität zugeführt werden.[265] Für einen Teil von Ansprüchen wegen Flugannullierung (Art. 5 Fluggastrechte-VO) lässt sich dies ebenfalls festhalten. Zwar besteht dort die Möglichkeit der Fluggesellschaft, außergewöhnliche Umstände gegen einen Anspruch ins Feld zu führen, es lassen sich jedoch auch standardisiert

[262] Vgl. *Meyer*, ZRP 2018, 233, 233, 236; siehe auch *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 117f.; s. a. *Rühl*, JZ 2020, 809, 816, die auch die Vermeidung von Fehlern sowie die Vereinheitlichung der Rechtsordnung und die Einhaltung des Gleichheitssatzes nennt.
[263] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 173; den (erprobenden) Einsatz von KI im Rahmen des Kostenfestsetzungsverfahrens ausdrücklich vorschlagend: Thesenpapier der Arbeitsgruppe „Modernisierung des Zivilprozesses“ im Auftrag der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://www.justiz.bayern.de/media/images/behoerden-und-gerichte/oberlandesgerichte/nuernberg/thesenpapier_der_arbeitsgruppe.pdf (Abruf am 26.02.2026, 13:01 Uhr).
[264] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 33f., vgl. auch *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 393f.; *Horn*, ZZP 136 (2023), 459, 462, 482ff.; *Wolff*, Algorithmen als Richter, 1. Aufl. 2022, S. 235.
[265] Siehe zudem die Risikoeinordnung von Fluggastrechteverfahren anhand des Anwendungsbeispiels MAKI als niedrig bis mittel, (jedoch) im Rahmen des entscheidungs*unterstützenden* Einsatzes informationstechnischer Systeme, den Abschnitt C.I.4.j.bb.

außergewöhnliche Umstände vordefinieren und automatisiert behandeln. Umstände, die nicht vordefiniert, aber gleichwohl unter den unbestimmten Rechtsbegriff der außergewöhnlichen Umstände zu subsumieren wären, könnten dann zwar nicht vom System selbst angemessen verarbeitet werden, hier wäre allerdings eine automatische Aussteuerung an den menschlichen Richter gestaltbar, sodass zumindest für die klar vordefinierten Fälle eine automatisierte Behandlung erfolgen könnte und die rechtlichen Rahmenbedingungen ebenfalls für komplexere Konstellationen über die menschliche Entscheidungsfindung sichergestellt wären.[266]
Auch durch die sukzessive Ausweitung des KI-Einsatzes für verschiedene Verfahrenskategorien kann perspektivisch ein zunehmend engmaschigeres Netz erreicht werden, das spürbar für die Justiz erhebliche Entlastung in standardisiert behandelbaren, einfachen Fall- und Rechtsstreitigkeiten bringt.
Bei der Entwicklung der jeweiligen Systeme sollte soweit möglich auf bestehende Systeme zurückgegriffen und somit **Synergiepotential** genutzt werden, wie dies beispielsweise im Rahmen des Projekts MAKI, das neben den zivilrechtlichen Einsatzszenarien z. B. auch im Asylverfahren zur Anwendung kommen soll, der Fall ist. Sofern möglich, sollten Systemfunktionen daher modular erstellt werden, insbesondere kann eine grundsätzlich explizit programmierte (regelbasierte) Software an verschiedenen Stellen Systeme maschinellen Lernens einbinden, die Teilaspekte (für die sie spezialisiert sind und insoweit auch die notwendige Qualität aufweisen) bearbeiten.[267] Im Rahmen der Aufsicht und Kontrolle kann ein derartiges Vorgehen bei auftretenden Fehlfunktionen ggf. das schnellere Auffinden des Fehlers samt minimalinvasiver Korrektur im Code ermöglichen.

(2) Verfahrensrechtliche Ausgestaltung des entscheidungsersetzenden Einsatzes

Anknüpfend daran kann der entscheidungsersetzende Einsatz informationstechnischer Systeme in verschiedenen verfahrensrechtlichen Ausgestaltungen erfolgen und erprobt werden. Einerseits kann das System im Rahmen einer **automatisch entscheidenden Instanz** (z. B. „0.-Instanz“) eine *letztverbindliche Entscheidung* – einem Urteil oder Beschluss eines menschlichen Richters entsprechend – treffen (sofern ein *hinreichend einfach gelagerter Fall* vorliegt). Hiermit könnte ein System zur effizienten Geltendmachung subjektiver Rechte etabliert werden, das schnell sowie effektiv zu einem Titel führt und eine Entlastung der Gerichte erreicht.[268] (Zu einer weitergehenden Betrachtung siehe auch den Abschnitt zur „Verfügbarkeit von Rechtsbehelfen“, C.II.7)
Eine weitere Option ist, das **System dem Mahnverfahren gem. §§ 688–703d ZPO ähnlich** auszugestalten, indem das System eine zwar verbindliche Entscheidung trifft, die jedoch unter der (alleinigen) Voraussetzung eines (ggf. fristgebundenen) Antrags

[266] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 90.
[267] *Horn*, ZZP 136 (2023), 459, 462, 482ff.; *Wolff*, Algorithmen als Richter, 1. Aufl. 2022, S. 235; siehe auch indes in Bezug auf entscheidungsunterstützende informationstechnische Systeme Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 125; vgl. auch *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 22f.
[268] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 116f., 174f.

eines Verfahrensbeteiligten einer menschlichen, richterlichen Kontrolle unterzogen wird.[269] Ein solches Verfahren würde den (materiellen) Rechtsprechungsbegriff nicht erfüllen, damit jenseits der Erfordernisse des Art. 92 GG agieren, und einen breiteren rechtlichen Anwendungsbereich für automatisierte Verfahren eröffnen, insbesondere auch für komplexere Sachverhalte und Streitigkeiten.[270] Mit Blick auf Art. 20 Abs. 3 GG ist auch bei einer derartigen Ausgestaltung im Ergebnis die Gesetzesbindung sicherzustellen. Dies kann, wie bereits derzeit im Mahnverfahren oder im Urkundenprozess erfolgend, über Rechtsbehelfsmöglichkeiten geschehen (siehe zur Thematik der Verfügbarkeit von Rechtsbehelfen weitergehend Abschnitt C.II.7).
Insgesamt und auch bei der Ausgestaltung als Mahnverfahrensäquivalent sollte jedoch eine hinreichend **hohe Quote an Genauigkeit** erreicht werden; im ähnlichen Zusammenhang werden für die Praxistauglichkeit für Einzelausgaben von Systemen maschinellen Lernens anhand des Bewertungsmodells F1-Score (errechnet aus dem harmonischen Mittel von Präzision [der Anteil an korrekt positiv klassifizierten Werten im Verhältnis zu allen – auch fälschlich – positiv klassifizierten Werten] und Sensitivität [dies misst, wie gut das Modell tatsächlich positive Fälle erkennt][271]) Zielwerte von 95 % vorgeschlagen, indes mit Hinweis darauf, dass genaue, praxisnutzbare Werte noch konkreter zu bestimmen und insgesamt ein verhältnismäßig hoher Wert an Genauigkeit zu verfolgen ist.[272] Das Qualitätsniveau sollte dem Kontext entsprechend präzisiert werden, auch sollten Prozesse eingerichtet werden, mittels derer der zu gewährleistende Stand der Technik ausgearbeitet wird.[273] Es sollte eine Verständigung über die konkreten Qualitätsmaßstäbe erfolgen, wobei Genauigkeit und Innovationshemmnisse – jedoch mit gezieltem Blick auf den judikativen Einsatz – im Auge behalten werden sollten; auch Forschungsförderung sollte in diesem Zusammenhang als Mittel in Betracht gezogen werden.[274] Im gleichen Zug sollten auch effektive Sanktionierungsmaßnahmen sowie Haftungsfragen in die Betrachtung einbezogen werden.[275] So könnte bei Fehlfunktionen eine staatliche und/oder private Haftung ggf. des privaten Herstellers des informationstechnischen Systems in Betracht genommen werden.[276] Bei fehlender Genauigkeit erscheint die Akzeptanz seitens der Bürgerinnen und Bürger fraglich und der gewünschte Effizienzgewinn könnte leiden. Geringere Kosten – wie beim derzeitigen Mahnverfahren – im Vergleich zu einem regulären gerichtlichen Verfahren, können als Anreiz zur Nutzung gesehen werden.[277] Vor dem Hintergrund dieser Aspekte ist **zu empfehlen,** einen entscheidungsersetzenden Einsatz aufgrund der derzeitigen Leistungsfähigkeit der Systeme v. a. in nur einfachen Fallkonstellationen vorzusehen. Hierbei sollten regelbasierte Systeme eingesetzt werden und dies im Sinne einer echten Instanz, um nachhaltige Effizienzgewinne zu bewirken. Angesichts der derzeit noch fehlenden

[269] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 116f.
[270] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 174ff.
[271] Siehe zur Sensitivität, die dieser als „Belastbarkeit“ beschreibt, weiter vertiefend *Quack*, Automation zivilrechtlicher Entscheidungsfindung, 1. Aufl. 2025, S. 233 (insb. Fn. 158). Die Sensitivität kann auch wie folgt beschrieben werden: Verhältnis: Systemseitig richtigerweise gegebene positive Bewertung im Verhältnis zu tatsächlich enthaltenen positiven Fällen im Datensatz.
[272] *Quack*, Automation zivilrechtlicher Entscheidungsfindung, 1. Aufl. 2025, S. 233ff.; *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 30.
[273] *Wischmeyer*, AöR 143 (2018), 1, 25.
[274] *Wischmeyer*, AöR 143 (2018), 1, 25.
[275] Vgl. *Wischmeyer*, AöR 143 (2018), 1, 25f.; vgl. auch *Turbék*, Justiz und KI, Friedrich-Naumann-Stiftung für die Freiheit (Hrsg.), 2026, abrufbar unter https://shop.freiheit.org/#!/Publikation/2184 (Abruf am 09.10.2026, 15:40 Uhr), S. 33.
[276] Vgl. zur Arzthaftung im Kontext des KI-Einsatzes *Meyer*, ZRP 2018, 233, 236.
[277] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 175f.

Verlässlichkeit von Systemen maschinellen Lernens ist von dem Einsatz im entscheidungsersetzenden Bereich (auch bei mahnverfahrensgleicher Ausgestaltung) eher abzuraten, um negative Effekte, insbesondere mit Blick auf die Akzeptanz der Rechtssuchenden und mit Blick auf das Vertrauen der Bürgerinnen und Bürger in das Funktionieren des Rechtsstaates zu vermeiden.[278] Letztlich steht ihr Einsatz im entscheidungsersetzenden Bereich derzeit mit der Gesetzesbindung, Art. 20 Abs. 3 GG, in Konflikt, da die Einhaltung dieser derzeit systemseitig nicht sichergestellt werden kann.[279]

(3) Jederzeitige Aussteuerungs(möglichkeit) in komplexeren Fallkonstellationen

Die für die Entscheidung des informationstechnischen Systems relevanten Tatsachen können über Abfrage- und Eingabemasken strukturiert erhoben werden, z. B. über klar definierte Online-Formulare. Ausgehend vom Recht auf rechtliches Gehör muss aber zusätzlich die freie Äußerung, unabhängig von einer solchen (starren) Abfrage- und Eingabemaske, möglich sein.[280] Denkbar ist die Vorgabe, konkrete Begriffe zu verwenden oder konkrete, standardisierte Eingabemöglichkeiten – insbesondere für regelmäßig wiederkehrende und gut standardisierbare Punkte – zur Verfügung zu stellen (z. B. anklickbare Felder, etwa für das Anerkenntnis nach § 308 S. 1 ZPO eines bestimmten Antrags).[281] Dennoch ist stets vorstellbar, dass bestimmte Aspekte eines Rechtsstreits nicht formularmäßig im Vorfeld abdeckbar sind. Daher sollte systemseitig eine **Aussteuerungsmöglichkeit** an einen menschlichen Richter bzw. eine menschliche Richterin bei komplexeren bzw. atypischen Gestaltungen vorgesehen werden, die das System nicht vollständig erfassen kann.[282]
Für die **Informationsvermittlung** gegenüber den Parteien – wenn diese ein Informationsgesuch an das Gericht stellen – kommt für den Fall des Einsatzes eines entscheidungsersetzenden informationstechnischen Systems ein gestuftes Vorgehen bzw. eine **gestufte Architektur** in Betracht: Die Informationsvermittlung bezogen auf den prozessualen Kontext kann seitens eines entscheidungsersetzenden informationstechnischen Systems in einfach gelagerten Fällen, bei denen diese einem Automatismus gleicht, durch das System selbst, erfolgen, indem es die Daten und Informationen aus der Akte an die Parteien leitet.[283] Abwägungsentscheidungen zwischen (etwaigen) Geheimhaltungsinteressen und dem Informationsgesuch einer Partei vermag ein informationstechnisches System derzeit nicht zu leisten.[284] In diesen Fällen ist das Handeln menschlicher Richter bzw. Richterinnen gefragt.[285] Soweit einfach gelagerte Fälle von Informationsgesuchen ohne Abwägungsentscheidung erfasst werden können, sollten diese bereits systemseitig berücksichtigt werden können. Dies erscheint im Interesse einer Arbeitsentlastung der Gerichte sowie eines schnellen Nachkommens der Informationsgesuche vorteilhaft. Hingegen sollten Fälle, in denen Informationsgesuche (komplexeren) Abwägungsentscheidungen begegnen (zumindest derzeit) an menschliche Richterinnen und Richter ausgesteuert werden.

[278] Siehe in diesem Zusammenhang jüngst *Choi*, Off-The-Shelf Large Language Models Are Unreliable Judges, Journal of Empirical Legal Studies 2026, 1ff.
[279] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 146.
[280] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 229ff., vgl. auch S. 334f., 344.
[281] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 335.
[282] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 230ff.
[283] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 232.
[284] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 232.
[285] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 232.

Beim entscheidungsersetzenden Einsatz *regelbasierter* Systeme sollte bereits im Design die Aussteuerung angelegt sein; es sollte eine Systemgestaltung verfolgt werden, dass das **Vorliegen unbekannter Konstellationen** im System bereits vorgedacht wird und bei Vorliegen eines solchen unbekannten Falls – den das System nicht selbst als vordefiniert verarbeiten kann – eine **automatisierte Aussteuerung zu einem menschlichen Richter bzw. einer menschlichen Richterin erfolgt**.[286] So kann nämlich beispielsweise eine Pflicht zur mündlichen Verhandlung entstehen, wenn der persönliche Eindruck des Betroffenen nötig ist.[287] Freilich sollten „typische“ außergewöhnliche Umstände und erst recht regelmäßig auftretende Sachverhaltskonstellationen – soweit möglich – vordefiniert und im System zur Verarbeitung implementiert werden.[288]
Ebenso sollte eine **Aussteuerung** des Falls **an einen menschlichen Richter bzw. eine menschliche Richterin auf Antrag** vorgesehen werden. Hierbei sollten Mechanismen, wie eine hinreichende Substantiierung (z. B. dass kein hinreichend einfach gelagerter Fall vorliegt), vorgesehen werden, um einer „missbräuchlichen“ Aussteuerung, etwa zur Verfahrensverzögerung, vorzubeugen.[289] Insgesamt ist hierbei ein Ausgleich insbesondere zwischen Praktikabilität, Zugänglichkeit und Missbrauchsverhinderung anzustreben.[290] Hinsichtlich der Aussteuerung sollte sowohl bereits bei der Eingabe der Daten als auch bei der Aussteuerung selbst ein strukturierter Parteivortrag vorgesehen werden; die Ansätze zu einem strukturierten Parteivortrag sind vielfach bereits historisch aufgekommen und im Rahmen der Digitalisierung erneut stark in den Blick geraten.[291] Der ausgesteuerte Fall sollte dem menschlichen Richter möglichst benutzerfreundlich und strukturiert aufbereitet zur Verfügung gestellt werden.

(4) Vermeidung systemseitiger Halluzinationen

Regelbasierte Systeme arbeiten deterministisch, bei identischer Eingabe folgt eine identische Ausgabe.[292] Halluzinationen sind insoweit primär eine Limitation informationstechnischer Systeme maschinellen Lernens.[293] Insoweit sollte zunächst ein weitgehender Einsatz regelbasierter Systeme verfolgt werden. Soll darüber hinaus auch die Einbindung informationstechnischer Systeme maschinellen Lernens forciert werden, kann in diesem Zusammenhang eine Grundstruktur anhand eines regelbasierten Systems in Betracht kommen, in dem an geeigneten Stellen punktuell informationstechnische Systeme maschinellen Lernens hohen F1-Scores bausteinartig mit Teilbereichen befasst sind, deren Ergebnisse in eine Gesamtentscheidung gefügt werden.[294]
Mit Blick auf die systemseitige Verringerung der Halluzinationsquoten kann ergänzend auf die Ausführungen in den Handlungsoptionen und -empfehlungen für den entscheidungsunterstützenden Einsatz informationstechnischer Systeme in diesem

[286] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 90, 145.
[287] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 226 m. w. N.
[288] Vgl. *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 145.
[289] Vgl. *Horn*, ZZP 136 (2023), 459, 486.
[290] Vgl. *Horn*, ZZP 136 (2023), 459, 486.
[291] Vgl. *Mielke/Wolff*, Der strukturierte Parteivortrag im Zivilprozess, in: Schweighofer/Saarenpää/Eder/Zanol/Schmautzer/Kummer/Hanke (Hrsg.), Recht DIGITAL – 25 Jahre IRIS, 2022, S. 200ff.
[292] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 21f.
[293] Vgl. nur *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 151f.
[294] *Quack*, Automation zivilrechtlicher Entscheidungsfindung, 1. Aufl. 2025, S. 222ff., 232ff., 235.

Kapitel verwiesen werden (siehe Abschnitt C.II.2.c.bb(1)). In Anbetracht der rapiden Fortentwicklung im Bereich der Systeme maschinellen Lernens und daraus ggf. folgenden Techniken zur Vermeidung systemseitiger Halluzinationen sollte der technische Fortschritt eng begleitet werden.
Schlussendlich ist derzeit aufgrund der fehlenden Verlässlichkeit im Bereich des entscheidungsersetzenden Einsatzes allerdings eher von der Verwendung von Systemen maschinellen Lernens abzuraten. Auch aktuell zeigt sich die fehlende Verlässlichkeit immer wieder. So ist z. B. erst kürzlich bekannt geworden, dass informationstechnische Systeme maschinellen Lernens (konkret ein KI-Modell des für die KI ChatGPT bekannten Unternehmens OpenAI) versucht haben, selbst erstellte Dateien ins Internet hochzuladen, um sie als Quellen heranziehen zu können.[295] Dem Einsatz entscheidungsersetzender informationstechnischer Systeme maschinellen Lernens steht die Gesetzesbindung entgegen, deren systemseitige Einhaltung derzeit nicht sichergestellt werden kann.[296]

cc. Zusammenfassung

Ein entscheidungsersetzender Einsatz ist derzeit nur in hinreichend einfach gelagerten Fällen zulässig. Hierunter zu verstehen sind Fälle mit überschaubaren Tatsachen- und Rechtsfragen sowie einem geringen Auslegungsspielraum, weswegen der Weg, auf dem das informationstechnische System zu seiner Entscheidung gelangt, einer mathematisch-logischen Rechenoperation gleichkommt. Insbesondere bei komplexen Abwägungsentscheidungen scheidet der entscheidungsersetzende Einsatz aus. Daher muss jederzeit eine systemseitige Aussteuerung an einen menschlichen Richter bzw. eine menschliche Richterin, auch auf Antrag, möglich sein, sobald sich zeigt, dass ein solcher Fall nicht vorliegt. Halluzinationen sind zu vermeiden. Ist dies nicht sichergestellt, ist ein entscheidungsersetzender Einsatz unzulässig.
Derzeit gewährleisten nur regelbasierte Systeme die Einhaltung der aufgezeigten Grenzen. Hervorzuheben ist, dass die Nichtnutzung informationstechnischer Systeme mit Kosten verbunden ist, da Entlastungspotentiale ungenutzt bleiben. Erste und auch notwendige Handlungsoption ist daher die **Identifikation relevanter und möglicher Einsatzbereiche**, z. B. Massenverfahren wie Fluggastrechteverfahren, Verfahren nach dem VDuG, die Kostenfestsetzung nach §§ 103, 104 ZPO sowie die Abwicklung von Verkehrsunfällen. Diese sollten durch interdisziplinär besetzte, Richterinnen und Richter einschließende, Arbeitsgruppen ermittelt und die Entscheidungskriterien in einem Entscheidungsbaum präzise herausgearbeitet werden. Systeme können zunächst für Teilbereiche entwickelt und sodann modular erweitert werden. Dabei sollte zur Nutzung von Synergien auf bestehende Systeme zurückgegriffen werden.
Verfahrensrechtlich kommt eine „0.-Instanz“ mit verbindlicher gerichtlicher Entscheidung oder eine dem Mahnverfahren ähnliche Ausgestaltung in Betracht. Die letztverbindliche Entscheidung würde bei letzterem auf Antrag einem Richter oder einer Richterin obliegen. In beiden Fällen sollte eine hinreichend hohe Genauigkeit erreicht werden. Qualitätsmaßstäbe sowie Haftungs- und Sanktionsfragen sind zu klären. **Empfohlen wird der entscheidungsersetzende Einsatz mittels regelbasierter Systeme**, angesichts des derzeitigen Stands der Technik, und hierbei eine Ausgestaltung im Sinne einer **echten Instanz**, um nachhaltige Effizienzgewinne

[295] *Sieb*, OpenAI veröffentlicht weitere KI-Zwischenfälle, abrufbar unter https://www.tagesschau.de/ausland/amerika/openai-kuenstliche-intelligenz-100.html (Abruf am 26.09.2026, 22:00 Uhr).
[296] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 146.

zu bewirken. Von dem Einsatz von Systemen maschinellen Lernens ist – auch bei einer Ausgestaltung ähnlich dem Mahnverfahren – im entscheidungsersetzenden Bereich derzeit eher abzuraten, um negative Auswirkungen auf die Akzeptanz der Rechtsschutzsuchenden und das Vertrauen in das Funktionieren des Rechtsstaates zu vermeiden; ihr Einsatz stünde zudem ohnehin derzeit mit der Gesetzesbindung nach Art. 20 Abs. 3 GG in Konflikt, deren Einhaltung systemseitig derzeit nicht sichergestellt werden kann.
Die entscheidungsrelevanten Tatsachen können über strukturierte Abfrage- und Eingabemasken erhoben werden. Eine freie Äußerung muss im Hinblick auf das rechtliche Gehör hierbei jedoch gewährleistet bleiben. Eine Aussteuerung von atypischen, nicht vom System erfassten Konstellationen sowie von Informationsgesuchen, die eine menschliche Abwägung erfordern, sollte im Systemdesign angelegt sein und automatisiert erfolgen. Typische Sachverhalte bzw. Sachverhalte ohne Abwägungsnotwendigkeit sollten hingegen soweit möglich vordefiniert und systemseitig bearbeitet werden. Für die Aussteuerung auf Antrag sollten Maßnahmen zur Missbrauchsverhinderung vorgesehen werden, wie eine hinreichende Substantiierung des Aussteuerungsantrags, wobei insgesamt ein Ausgleich zwischen Praktikabilität, Zugänglichkeit und Missbrauchsverhinderung angestrebt werden sollte. Der ausgesteuerte Fall sollte dem zuständigen Richter oder der zuständigen Richterin strukturiert aufbereitet zur Verfügung gestellt werden.
Zur Vermeidung von Halluzinationen sollte vorrangig auf deterministisch arbeitende regelbasierte Systeme gesetzt werden. Sollte der Einsatz von Systemen maschinellen Lernens forciert werden, kommt der Einsatz von Systemen maschinellen Lernens mit einem hohen F1-Score bausteinartig innerhalb einer regelbasierten Grundstruktur in Betracht; schlussendlich ist von ihrem Einsatz jedoch derzeit mangels Sicherstellung des Einhaltens der Gesetzesbindung abzuraten. Der voranschreitende technische Fortschritt ist eng zu begleiten.

c. Rote Linien für einen entscheidungsunterstützenden Einsatz sowie Handlungsleitlinien

Neben den übergreifenden roten Linien, die unter C.II.2.a bereits eingehend betrachtet worden sind, ergeben sich für den spezifischen Einsatz als entscheidungs*unterstützendes* System weitere rote Linien.

aa. Rote Linien entscheidungsunterstützender Einsatz

Auch im Hinblick auf den entscheidungsunterstützenden Einsatz ist festzustellen: Aufgrund der gegenwärtigen technischen Limitierungen der Leistungsfähigkeit informationstechnischer Systeme können diese die Einhaltung der rechtlichen Anforderungen nur in hinreichend einfach gelagerten Fällen sicherstellen.[297] Nutzt der Richter oder die Richterin unterstützende, administrative oder inhaltliche Funktionen eines informationstechnischen Systems, muss er oder sie dessen Arbeitsergebnis kontrollieren und nach Bedarf korrigieren (**Zwingendes Erfordernis richterlicher Kontrolle und ggf. Korrektur**).[298] Ihm bzw. ihr müssen effektive Möglichkeiten zur **substanziellen Kontrolle und Korrektur** der von dem informationstechnischen

[297] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 87ff., 108ff., 154, 240f., 255f., 265ff.
[298] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 90f., 108ff., 119f., 154f., 177ff., 221ff., 237f., 240ff., 255f., 265ff., 335f., 358f.

System erzeugten Ergebnisse eingeräumt werden.[299] Der Richter oder die Richterin trägt letztlich die **Letztverantwortung** für die Gesetzeskonformität der Entscheidung. Intensität und Umfang der notwendigen Kontrollerfordernisse sind wiederum abhängig von den drei bereits genannten Faktoren: Komplexität des Anwendungsfeldes, Grad der Automatisierung und Funktionalität des Systems (administrativ vs. sachlich-inhaltlich).[300] Hierbei gilt, dass je höher der Automatisierungsgrad bzw. je erweiterter die Funktionalität ist, desto intensiver hat die Kontrolle des Richters oder der Richterin auszufallen.[301] Besondere Sorgfalt ist seitens der Richterschaft demnach an den Tag zu legen, wenn das informationstechnische System inhaltliche Aufgaben erfüllt.[302]
Die richterliche Letztverantwortung erstreckt sich dabei insbesondere dahingehend, dass alle verfahrensrelevanten Informationen zutreffend erfasst und verarbeitet wurden.[303] Der Richter oder die Richterin trägt insbesondere auch die Verantwortung dafür, dass sich ein systemseitiger *Bias* und systemseitige Halluzinationen nicht in der zu treffenden Endentscheidung und anderen verfahrensgestaltenden Maßnahmen niederschlagen.[304] Die Letztentscheidungsverantwortung kann durch psychologische Effekte, wie einen *Automation Bias*[305] oder *Default-Effekt*[306] beeinträchtigt sein.[307] Insofern sind bei einem entscheidungsunterstützenden Einsatz geeignete systemseitige oder sonstige Vorkehrungen zu treffen, um solche Effekte möglichst auszuschließen.[308] Richter und Richterinnen müssen letztlich über die Funktionsweise informationstechnischer Systeme und die ihnen innewohnenden Limitierungen entsprechend informiert werden, damit sichergestellt ist, dass sie die konkreten Möglichkeiten und Grenzen des Systems korrekt einschätzen und die Kontrollerfordernisse hinreichend wahrnehmen (können).[309]
Dies alles folgt insbesondere aus der Menschenwürdegarantie gemäß Art. 1 GG[310], dem Justizgewährungsanspruch gem. Art. 20 i. V. m. Art. 2 Abs.1 GG[311], dem Rechtsstaatsprinzip gemäß Art. 20 Abs. 3 GG[312], aus Art. 92, 97 GG[313], aus

[299] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 90f., 108ff., 119f., 154f., 177ff., 221ff., 237f., 240ff., 255f., 265ff., 335f., 358f.
[300] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 90f., 108ff., 119f., 154f., 168, 177ff., 240f., 255f., 265ff., 344f., 358f., vgl. auch S. 236.
[301] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 87f., 108ff., 154f., 168, 177f., 240f., 255f., 265ff., vgl. auch S. 236.
[302] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 90f., 108ff., 154f., 177ff., 221ff., 240f., 255f., 265ff.
[303] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 108ff., 119f., 255f., 265ff., 344f., 358f.
[304] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 90f., 108ff., 154f., 240ff., 265ff., 358f.
[305] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 90f., 108ff., 119f., 127f., 153f., 202ff., 221f., 240ff., 265ff., 358f.; zum *Automation Bias* selbst siehe S. 202ff., 211.
[306] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 202ff., 221f.; zum *Default-Effekt* selbst siehe S. 203f., 211.
[307] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 90f., 108ff., 119f., 127f., 153f., 202ff., 221f., 240ff., 265ff., 358f.; zu weiteren *Bias*, wie dem sog. *Confirmation Bias* oder dem sog. *Technical Bias* siehe S. 262f.
[308] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 119ff., 127f., 202ff., 221f., 242ff., 335f., 348f., 358f.
[309] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 154f.
[310] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 90f.
[311] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 119f.
[312] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 154f.
[313] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 154, 177ff., 202ff.

Art. 101 Abs. 1 Satz 2 GG,[314] aus Art. 6 Abs. 1 EMRK[315] und aus Art. 47 Abs. 1, Abs. 2 GRCh[316]. Auch Art. 14 KI-VO bestimmt Entsprechendes für Hochrisiko-KI-Systeme.

bb. Handlungsoptionen und -empfehlungen

(1) Substantielle Kontroll- und Korrekturmöglichkeiten

Zur Umsetzung der vorstehenden rechtlichen Rahmenbedingungen sind mehrere Handlungsoptionen denkbar. Richterinnen und Richter sollten ganz allgemein in die Lage versetzt werden, die wesentlichen Schritte des Entscheidungsprozesses des Systems, die die systemseitige Entscheidung tragen, zu rekonstruieren.[317] Insgesamt ist eine möglichst weitgehende Nachvollziehbarkeit zu befürworten, die sich auch auf Kriterien und Aspekte, anhand derer das System das Ergebnis generiert hat, erstreckt.[318]
Eine zu befürwortende Handlungsoption zur Erreichung dieses Ziels besteht beispielsweise darin, die richterliche Aufsicht und Kontrolle durch eine **herkunftssensitive Benutzeroberfläche** zu unterstützen. Das System sollte hierüber anzeigen, woher es die verarbeitete Information bezogen hat, damit der Richter bzw. die Richterin nachvollziehen kann, welche Informationen das informationstechnische System in dem erzeugten Ergebnis verarbeitet hat, und die Richtigkeit der Systemausgabe überprüft werden kann.[319] Dies kann im Detail auf unterschiedliche Weise erfolgen:
Erstens kann eine systemseitige Unterstützung über die **Art und Weise der Darstellung von verarbeiteten Informationen** erfolgen: So ist beispielsweise gleichsam möglich, verarbeitete Informationen im Originalwortlaut darzustellen oder den Inhalt der verarbeiteten Information in Form einer sinngemäßen Zusammenfassung bzw. Paraphrasierung anzugeben. Im Grundsatz dürfte – soweit dies möglich ist – aufgrund des geringeren Fehlerrisikos die **exakte Textübernahme aus der Originalquelle** vorzugswürdig sein. Die **Übernahme** von Informationen **in einer sinngemäßen Zusammenstellung oder Wertung** des Systems ist demgegenüber weniger vorzugswürdig, da sie fehleranfällig ist und unter Umständen Wertungen enthält, für deren Vornahme entsprechende Systeme weder trainiert noch geeignet sind.[320] Erstellt ein System daher beispielsweise eine Relationstabelle bzw. einen Tatbestand für ein Urteil, ist die Übernahme von Originaltext der sinngemäßen Zusammenfassung von Akteninhalt gegenüber vorzugswürdig, sofern nicht aufgrund

[314] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 221ff.
[315] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 240ff.
[316] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 265ff.
[317] *Martini/Botta/Nink/Kolain*, Automatisch erlaubt? Fünf Anwendungsfälle algorithmischer Systeme auf dem juristischen Prüfstand, 2020, S. 51f.; siehe auch *Köhler*, Richterliche Verantwortung und Künstliche Intelligenz, 1. Aufl. 2026, S. 207f.
[318] Vgl. *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 182; siehe auch Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 97ff., 125.
[319] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 179.
[320] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 233, 236.

der schieren Datenmenge eine Zusammenfassung zum Erhalt der Funktionalität des Systems angezeigt ist. Letzteres kann etwa bei sehr umfangreichen Schriftsatzdokumenten der Fall sein. Eine sinngemäße Zusammenfassung ist somit nicht gänzlich abzulehnen. Sie bringt aufgrund der enthaltenen Systemwertungen und der systembedingten Fehlermöglichkeiten allerdings weitergehende Kontrollerfordernisse mit sich, die beachtet werden müssen.

Zweitens kann die Gestaltung einer herkunftssensitiven Benutzeroberfläche die richterliche Aufsicht und Kontrolle über die **Möglichkeit** unterstützen**, die Originalquellen der verarbeiteten Informationen abzurufen und zum Zweck der Überprüfung der Richtigkeit der Systemergebnisse den Inhalt der Originalquellen einzusehen und zu durchsuchen:** Einer effektiven Kontrolle und Aufsicht durch menschliche Richterinnen und Richter ist es zuträglich, die **Originalquellen**[321] **möglichst niedrigschwellig** und medienbruchfrei[322] **abrufbar** zur Verfügung zu stellen und die **Möglichkeit einer intensiven Recherche in diesen Quellen** zu ermöglichen.

Nur so lassen sich Ausgabeergebnisse des Systems effektiv und effizient überprüfen,[323] denn das System mag in diesen Quellen relevante Informationen übersehen haben, deren Relevanz erst bei einer richterlichen Detailkontrolle offenbar wird. Auch mag es bei der Datenübernahme zu Fehlern kommen, etwa dann, wenn Dokumente der Akte erst in ein maschinenlesbares Datenformat (z. B. JSON) überführt werden müssen. Eine Prüfung der Originalquelle sollte somit in jedem Fall ermöglicht werden. Hierbei sollte es zum **Gold-Standard** gehören, dass eine **Anzeige der Originalquellen** direkt über die Benutzeroberfläche des Systems und mit Bezug zur verarbeiteten Information ermöglicht wird. Per Klick sollte man diese und weitere Akteninformationen im **Volltext** abrufen können. In dem referenzierten Dokument sollte die entsprechende Textpassage sofort systemseitig angesteuert[324] und hervorgehoben angezeigt werden (z. B. durch farbliche Kenntlichmachung bzw. Unterlegung). Auch kann dem eine zunächst Pop-up-artige Darstellung des Dokuments vorangehen. Sofern sich hieraus Korrekturbedarf ergibt, sollte eine **Korrektur** im Text **ebenso niederschwellig** ermöglicht werden. Zu vermeiden sind demgegenüber Systembrüche und die Erschwerung menschlicher Kontrolle durch ein notwendiges „Herumspringen bzw. -schalten“ zwischen unterschiedlichen Softwaresystemen (z. B. zwischen KI-Unterstützungssoftware und e-Akte). Idealerweise ist ein KI-Assistenzsystem in das e-Aktensystem integriert.[325]

Eine effektive Aufsicht und Kontrolle durch menschliche Richterinnen und Richter setzt ferner voraus, dass diese die zugrundeliegende Methodik und die möglichen Fehlerquellen des Systems einschätzen können.[326] Bei der Gestaltung von KI-

---

[321] *Schneider/Wundenberg/Sharaf*, ZHR 189 (2025), 259, 276f.

[322] Zur Begrifflichkeit des Medienbruchs siehe *Bundesministerium für Digitales und Staatsmodernisierung*, Lexikon, Stichwort „Medienbruch“, abrufbar unter https://www.digitale-verwaltung.de/Webs/DV/DE/aktuelles-service/faq-lexikon/glossar/functions/ozg-lexikon.html?lv3=20348760&lv2=20339810#doc20348760 (Abruf am 04.06.2026, 19:52 Uhr).

[323] Vgl. *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 115f.; vgl. *Schneider/Wundenberg/Sharaf*, ZHR 189 (2025), 259, 276f.

[324] Vgl. cum grano salis die Lesezeichen/Textmarken-Funktion bei Word.

[325] In diese Richtung *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 32f., wonach im Rahmen einer KI-Plattform der Justiz geeignete Schnittstellen definiert werden sollen, die eine nahtlose Einbindung von KI-Systemen in gängige e-Aktensysteme und Fachverfahren ermöglichen sollen.

[326] *Martini/Botta/Nink/Kolain*, Automatisch erlaubt? Fünf Anwendungsfälle algorithmischer Systeme auf dem juristischen Prüfstand, 2020, S. 52.

Unterstützungssystemen hat sich daher in der Entwicklung gegenwärtig die Praxis herausgebildet, **Unsicherheiten bei der Ergebnisaussage** durch das System selbst kenntlich machen zu lassen.[327] Dies geschieht über sog. **Scores**, die Schwellenwert-abhängig Einschätzungen generieren, mit denen das System signalisiert, wie „sicher“ bzw. „unsicher“ es sich mit der Richtigkeit der Ausgabe ist.[328] Mängel in der Datengrundlage bzw. in deren Verarbeitung sollen auf diese Weise ausgewiesen und ihre Auswirkung auf das systemseitige Ergebnis kenntlich gemacht werden. Erkennende Richter bzw. erkennende Richterinnen sollen hierüber bei Unsicherheiten zur ergänzenden Prüfung veranlasst werden.[329]
Entsprechende Scores verfolgen zwar einen guten Zweck, sie haben allerdings eine entscheidende Schwäche: Die ausgewiesenen Scores sind letztlich jeweils selbst Ausgaben, die das informationstechnische System anhand bestimmter, erlernter Muster berechnet hat. Sie sind daher ebenfalls nicht als verlässlich einzustufen. Liegt das System falsch, kann dies weitreichende Folgen haben. Bei einer Einschätzung einer Information als „unsicher“ wird sich die Richtigkeit einer Information zwar regelmäßig über eine erfolgte Prüfung des menschlichen Entscheiders verifizieren lassen, geht das System allerdings fälschlicherweise von einem hohen Sicherheitswert (z. B. „ganz sicher“) aus, so könnte dies den menschlichen Entscheider dazu verleiten, die Richtigkeit der Information nicht mehr zu überprüfen. Damit würde sich der entsprechende Fehler auch in der Entscheidung niederschlagen. Von der Verwendung derartiger Scores ist daher abzuraten.
Besonders die Leistungsfähigkeit informationstechnischer Systeme im Bereich der Verarbeitung und Strukturierung großer Datenmengen verspricht erhebliche Vorteile bei der schnellen Erledigung von Verfahren, die umfangreiche Datenmengen beinhalten (z. B. in Massenverfahren wie Dieselverfahren und in Verfahren bei Verletzungen von Regelungen der DSGVO, in denen Schriftsätze mehrere hundert Seiten umfassen).[330] Die Rationalisierungsgewinne, die der Einsatz entscheidungsunterstützender informationstechnischer Systeme bieten kann, hängen jedoch insbesondere davon ab, dass derartige Systeme hinreichend korrekt funktionieren. Der nutzende Richter bzw. die nutzende Richterin unterliegt insbesondere dem zwingenden Erfordernis der Kontrolle und ggf. Korrektur. Funktioniert ein entscheidungsunterstützendes informationstechnisches System daher unzureichend und muss der Richter bzw. die Richterin erheblichen Aufwand in die Kontrolle und Korrektur investieren, kann der Rationalisierungseffekt des Einsatzes im Vergleich zur herkömmlichen Arbeit ohne den Einsatz neutralisiert oder sogar ins Gegenteil verkehrt werden.[331] Anders als im Falle des entscheidungsersetzenden

[327] Die Ausgabe von Unsicherheiten befürwortend etwa *Dötterl*, RDi 2026, 414, 419.
[328] Vgl. *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 394ff.
[329] *Martini/Botta/Nink/Kolain*, Automatisch erlaubt? Fünf Anwendungsfälle algorithmischer Systeme auf dem juristischen Prüfstand, 2020, S. 52.
[330] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 115, vgl. auch die *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, Münchener Thesen zum Zivilprozess der Zukunft, abrufbar unter https://www.olg-duesseldorf.nrw.de/infos/Zivilprozess-der-Zukunft/Muenchener-Thesen-zum-Zivilprozess-der-Zukunft.pdf (Abruf am 26.02.2026, 18:27 Uhr), S. 9, These B.3; vgl. auch *Turbék*, Justiz und KI, Friedrich-Naumann-Stiftung für die Freiheit (Hrsg.), 2026, abrufbar unter https://shop.freiheit.org/#!/Publikation/2184 (Abruf am 09.10.2026, 15:40 Uhr), S. 11 m. w. N.
[331] Zudem können auf KI zurückzuführende Ungenauigkeiten negatives mediales Echo nach sich ziehen, was der Akzeptanz abträglich sein kann, vgl. *Zenthöfer*, Der peinliche Copy-Paste-Fehler hessischer Richter mit ChatGPT, abrufbar unter https://www.faz.net/aktuell/feuilleton/medien-und-film/medienpolitik/der-peinliche-copy-paste-fehler-hessischer-richter-mit-chatgpt-201105297.html (Abruf am 18.08.2026, 18:01 Uhr).

Einsatzes informationstechnischer Systeme bilden systemseitige Halluzinationen informationstechnischer Systeme maschinellen Lernens keine den Einsatz ausschließende rote Linie für den Einsatz des Systems (siehe hierzu Abschnitt C.II.2.b.aa), da der Einsatz der entscheidungsunterstützenden informationstechnischen Systeme mit dem zwingenden Erfordernis richterlicher Kontrolle und ggf. Korrektur einhergeht und einen menschlichen Richter bzw. eine menschliche Richterin die richterliche Letztverantwortung für die Entscheidung trifft.[332] Mit Blick insbesondere auf den Rationalisierungseffekt, der mit dem Einsatz entscheidungsunterstützender informationstechnischer Systeme einhergehen soll, sind Halluzinationen gleichwohl vor dem Hintergrund des Vorgenannten zu adressieren. Zur Vermeidung dieser sollten daher bereits systemseitig Gegenmaßnahmen etwa in Form von sog. **Retrieval-Augmented Generation (RAG)-Systemen** verwendet werden, bei denen Informationen zur Beantwortung der Anfrage aus Informationsquellen bezogen werden und sodann unter Zugrundelegung dieser Suchergebnisse das Ergebnis der Anfrage ausgegeben wird.[333] Hierbei kommen insbesondere justizinterne juristische Datenbanken in Betracht.[334] Es sollte jedoch beachtet werden, dass selbst größere Large Language Models (LLM), die regelmäßig Ausgaben generieren, die besser auf die zugrundeliegende Datenlage gestützt sind, substantiell (ca. 25 %) unbelegte Aussagen aufweisen; auch für den juristischen Gebrauch spezialisierte LLMs weisen nach aktuellen Tests Halluzinationsquoten zwischen 17 % und 33 % auf.[335] Optimierungen der Eingaben in das System können hierbei abmildernd wirken.[336] Auch bietet es sich an, Ausgaben in einzelne „atomare“ Informationen zu zerlegen und diese sodann gegen eine Wissensdatenbank validieren zu lassen; dies kann (zusätzlich) durch entsprechend spezialisierte informationstechnische Systeme geschehen.[337] Ergänzend kommen insbesondere systemische Konsistenzkontrollen der Ergebnisse in Betracht, da bei einer Abweichung voneinander bzw. bei einem Widerspruch untereinander mehrfach erzeugter Ausgaben dies auf eine Halluzination hindeutet.[338] Die Vermeidung von systemseitigen Halluzinationen ist mithin ein wesentlicher Baustein (auch [zu diesem Aspekt im Zusammenhang mit dem entscheidungsersetzenden Einsatz informationstechnischer Systeme siehe Abschnitt C.II.2.b.bb(4)]) für einen

---

[332] Zu Folgen von KI-Halluzinationen im Zivilprozess siehe *Riehm/Valerius*, RDi 2025, 612, 612ff.

[333] *Stolfo*, Groundedness in Retrieval-augmented Long-form Generation: An Empirical Study, in: Findings of the Association for Computational Linguistics: NAACL 2024, 2024, S. 1537f.; *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 115f.

[334] Vgl. *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 115.

[335] *Stolfo*, Groundedness in Retrieval-augmented Long-form Generation: An Empirical Study, in: Findings of the Association for Computational Linguistics: NAACL 2024, 2024, S. 1538; *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 115; *Turbék*, Justiz und KI, Friedrich-Naumann-Stiftung für die Freiheit (Hrsg.), 2026, abrufbar unter https://shop.freiheit.org/#!/Publikation/2184 (Abruf am 09.10.2026, 15:40 Uhr), S. 8, 17 führt unter Verweis auf mehrere Studien an, dass führende LLMs in 58 % bis 88 % der Fälle halluzinieren, in denen ihnen gängige juristische Aufgaben gestellt wurden.

[336] *Omar/Sorin/Collins/Reich/Freeman/Gavin/Charney/Stump/Bragazzi/Nadkarni/Klang*, Multi-model assurance analysis showing large language models are highly vulnerable to adversarial hallucination attacks during clinical decision support, Communications Medicine 5, 330 (2025), S. 1ff.; siehe zur Thematik der Verbesserung der Eingaben (u. a. des Prompt-Engineerings i. S. v. Regeln für gelungene Arbeitsanweisungen) an KI-Systeme *Kuhlmann*, NJW 2026, 2774, 2777f.

[337] *Min/Krishna/Lyu/Lewis/Yih/Koh/Iyyer/Zettlemoyer/Hajishirzi*, FACTSCORE: Fine-grained Atomic Evaluation of Factual Precision in Long Form Text Generation, Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing 2023, S. 12076f.

[338] *Manakul/Liusie/Gales*, SelfCheckGPT: Zero-Resource Black-Box Hallucination Detection for Generative Large Language Models, Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing 2023, S. 9004f.

(Rationalisierungsgewinne ermöglichenden) Einsatz entscheidungsunterstützender informationstechnischer Systeme.
Auch jenseits des entscheidungsersetzenden Einsatzes wird eine strukturierte Kommunikation – die auch den Parteivortrag (mit Unterteilung in Tatsachen- und Rechtsvortrag) betrifft – befürwortet.[339] Hierbei sollte die **Begrenzung des Vortragsumfangs** – KI-generierte Schriftsätze führen bereits jetzt, u. a. auch wegen ihres Umfangs und ihrer schnellen Generierung zu einer erheblichen Belastung von Gerichten[340] – möglich sein, was auch durch Kostenanreize unterstützt werden kann; hierbei werden frühe prozessleitende, gerichtliche Maßnahmen sowie Strukturierungstermine als zielführende Maßnahmen identifiziert.[341] Durch eine **strukturierte Datengrundlage** ist eine höhere Genauigkeit der Ergebnisse, die informationstechnische Systeme ausgeben, möglich, was sich freilich auch vorteilhaft im entscheidungsunterstützenden Einsatz niederschlagen kann.[342] Zugleich ermöglichen effiziente sowie strukturierte Parteivorträge eine **effizientere Aufsicht und Kontrolle der Arbeitsergebnisse**.[343] Idealerweise werden solche strukturierten Daten bereits zu Beginn des Verfahrens erhoben, so dass sich Plattformlösungen, die sowohl von der Justiz als auch von Bürgerinnen und Bürgern sowie von Rechtsanwältinnen und Rechtsanwälten genutzt werden, geradezu anbieten.[344]

(2) Kontrollmöglichkeiten und systemseitige Kontrollpflichten

Die Sicherstellung hinreichender richterlicher Kontrolle der Arbeitsergebnisse ist für die Einhaltung der rechtlichen Rahmenbedingungen für den legalen Einsatz von KI im

---

[339] Die *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, Münchener Thesen zum Zivilprozess der Zukunft, abrufbar unter https://www.olg-duesseldorf.nrw.de/infos/Zivilprozess-der-Zukunft/Muenchener-Thesen-zum-Zivilprozess-der-Zukunft.pdf (Abruf am 26.02.2026, 18:27 Uhr), S. 8f., These B.2; siehe auch *Dötterl*, RDi 2026, 414, 420.

[340] Siehe hierzu *dpa/xp/LTO-Redaktion*, „Wir stehen mit dem Rücken zur Wand“, abrufbar unter https://www.lto.de/recht/justiz/j/nordrhein-westfalen-sozialgerichte-wegen-ki-schriftsaetzen-ueberlastet (Abruf am 18.08.2026, 17:45 Uhr).

[341] Die *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, Münchener Thesen zum Zivilprozess der Zukunft, abrufbar unter https://www.olg-duesseldorf.nrw.de/infos/Zivilprozess-der-Zukunft/Muenchener-Thesen-zum-Zivilprozess-der-Zukunft.pdf (Abruf am 26.02.2026, 18:27 Uhr), S. 8f., These B.2; siehe auch Dötterl, RDi 2026, 414, 420.

[342] Vgl. *Sundararajan/Sripada/Reiter*, Input Matters: Evaluating Input Structure’s Impact on LLM Summaries of Sports Play-by-Play, Proceedings of the 18th International Natural Language Generation Conference 2025, S. 795ff.; *Schneider/Wundenberg/Sharaf*, ZHR 189 (2025), 259, 276.

[343] Vgl. hierzu auf Grundlage des sog. digitalen Basisdokuments *Mielke/Wolff*, RDi 2024, 361, 367f., siehe aber auch *Mielke/Wolff*, RDi 2024, 361, 370.

[344] Siehe zu entsprechenden Vorschlägen die Reformkommission Zivilprozess der Zukunft, insbesondere C.IV bis C.VI, *Reformkommission Zivilprozess der Zukunft*, Abschlussbericht der Reformkommission Zivilprozess der Zukunft im Auftrag des 3. Digitalgipfels der Justizministerinnen und Justizminister des Bundes und der Länder, abrufbar unter https://www.bmjv.de/SharedDocs/Publikationen/DE/Fachpublikationen/2025_Abschlussbereicht_Reformkomission_Zivilprozess.html (Abruf am 20.04.2026, 21:09 Uhr), S. 47ff., 53ff., 63ff.; siehe auch Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 111ff.

entscheidungs*unterstützenden* Bereich elementar.[345] Eine **Kontrollpflicht** folgt bereits aus der Gesetzesbindung, Art. 20 Abs. 3 GG.[346] Der Umfang der erforderlichen Kontrollintensität richtet sich letztlich an den drei bereits zuvor genannten Faktoren aus: Komplexität des Anwendungsfeldes, Grad der Automatisierung und Funktionalität des Systems (administrativ vs. sachlich-inhaltlich).[347] Die **Bandbreite der Handlungsoptionen** reicht letztlich von der bloßen Schaffung niederschwelliger Kontrollmöglichkeiten, die Richterinnen und Richter eigenverantwortlich wahrnehmen können, bis hin zu systemseitig vorgesehenen, notwendigen Kontrollschritten, die vor dem Signieren der finalen Entscheidung durchlaufen werden müssen. Auch der **Einbau bewusster Fehler**, die die Aufmerksamkeit und Fokussierung von Richterinnen und Richtern sicherstellen, ist **prinzipiell denkbar**, sofern sie nicht in die finale Entscheidung einfließen, sondern dem menschlichen Entscheider lediglich das Erfordernis von konkreter Aufmerksamkeit aufzeigen. Hierbei ist zu beachten, dass zwar Richterinnen und Richter bereits aufgrund ihrer hohen intrinsischen Motivation zur pflichtgemäßen Aufgabenerfüllung in den meisten Fällen Arbeitsergebnisse von informationstechnischen Systemen gewissenhaft überprüfen und die ordnungsgemäße und richtige Wiedergabe des Streitstoffs durch das System kontrollieren werden. Gleichzeitig sind auch Richterinnen und Richter in anderen Bereichen bereits fundiert nachgewiesenen psychologischen Effekten ausgesetzt, die eine Systemkontrolle unbewusst hindern bzw. erschweren können. Hierzu zählen Effekte wie der *Automation Bias* oder auch der *Default-Effekt*.[348] Diese Effekte gilt es auch bei der Diskussion von Handlungsoptionen und -empfehlungen zu berücksichtigen.
Ganz allgemein bietet es sich an, die Gestaltung von Kontrollmöglichkeiten und -pflichten anhand der drei die Kontrollanforderungen beeinflussenden Faktoren auszurichten. So ist zu empfehlen, z. B. ein **automationsgradabhängiges, funktionsabhängiges und komplexitätsabhängiges Kontrollmodell** vorzusehen. Bei rein administrativen Unterstützungsleistungen wie der Metadatenextraktion können beispielsweise bereits niedrigschwelligere Prüf- und Korrekturmöglichkeiten genügen. Extrahiert ein System Metadaten fehlerhaft und bezeichnet beispielsweise eine Klageschrift fälschlich als Klageerwiderung bzw. werden die Parteidaten fehlerhaft extrahiert, wird der erkennenden Richterin bzw. dem erkennenden Richter bei der Bearbeitung des Falls recht leicht auffallen, dass das Ergebnis fehlerhaft ist. Eine Korrektur wird durch eine einfache Umbenennung der Dokumente bzw. der Parteinamen bewerkstelligt werden können. Sofern das System entsprechende Korrekturmöglichkeiten vorsieht, wirft die Kontrolle und die Fehlerkorrektur keinerlei Schwierigkeiten auf. Je stärker das System allerdings inhaltsbezogene Tätigkeiten wahrnimmt und je höher der Automatisierungsgrad ist, desto enger sollten Kontrollintensitätsschritte ausgestaltet sein. Bei einem System, das automatisiert

[345] Siehe auch *Gsell/Pfündl*, RDi 2026, 457, 459, 461, die ein Verhalten fordern, das ausschließt, dass KI-Vorschläge ohne hinreichende Prüfung, Reflexion und Verantwortung durch den zuständigen Richter übernommen werden und auf die Besonderheiten bei der Kontrolle der Ergebnisse entscheidungsunterstützender informationstechnischer Systeme hinweisen; die Wichtigkeit umfassender Kontrolle der Inhalte anführend *Turbék*, Justiz und KI, Friedrich-Naumann-Stiftung für die Freiheit (Hrsg.), 2026, abrufbar unter https://shop.freiheit.org/#!/Publikation/2184 (Abruf am 09.10.2026, 15:40 Uhr), S. 31f.
[346] Die Notwendigkeit einer hinreichenden Prüfung stößt auf breiten Konsens in der Literatur, siehe nur *Gsell/Pfündl*, RDi 2026, 457, 462f. m. w. N.
[347] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 90f., 108ff., 119f., 154f., 168, 177ff., 240f., 255f., 265ff., 344f., 358f., vgl. auch S. 236.
[348] Hierzu *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 119f., 242ff., 265ff.; siehe auch *Dötterl*, RDi 2026, 414, 417f.

Akteninhalt in bereits vorformulierte Textbausteine einfügt, ist ein mittlerer Automatisierungsgrad gegeben, daher ergeben sich auch mittlere Kontrollerfordernisse. Unterbreitet das System aber beispielsweise aus dem Akteninhalt automatisiert und generativ eigene Textvorschläge, so ist aufgrund der Fehleranfälligkeit derzeitiger Systeme das höchste Kontrollerfordernis gegeben. Für weitreichendere Funktionen (mittleres und höchstes Kontrollerfordernis) könnte daher eine zwingende bzw. ausdrückliche Freigabe von (einzelnen) Ausgabeteilen vorgesehen werden. So ist beispielsweise denkbar und durchaus zu empfehlen, dass Richterinnen und Richter **vor Freigabe der Signaturmöglichkeit zunächst einen bestimmten, stichprobenmäßigen Anteil (z. B. 25–33 %) der verarbeiteten Systemergebnisse** (z. B. nach Zufall ausgewählt) **überprüfen müssen**. Dies entspricht beispielsweise gängiger Praxis bei in der Medizin genutzten, KI-basierten Unterstützungssystemen, die ebenfalls das Durchlaufen gewisser Prüfschritte voraussetzen, bevor die Behandlung (z. B. eine auf KI-gestützter Diagnostik beruhende Bestrahlung in der Krebstherapie) durch den behandelnden Arzt bzw. die behandelnde Ärztin freigegeben werden kann. Dies würde zwar nicht die nach Art. 20 Abs. 3 GG notwendige vollständige Prüfung der Systemergebnisse ersetzen, allerdings bestimmten Effekten, wie einem Technikübervertrauen, einen Gegenimpuls entgegenstellen. Selbstverständlich müsste die Überprüfungsmöglichkeit so niederschwellig und effektiv sein, dass sie keine unnötigen praktischen Hürden aufbaut und gleichzeitig eine tatsächliche inhaltliche Kontrolle sicherstellt. Für sich alleinstehend nicht geeignet sind beispielsweise Gestaltungen, wie sie aus dem Bereich des Einbezugs allgemeiner Geschäftsbedingungen bekannt sind, z. B. das bloße Setzen eines Häkchens bzw. das bloße Herunterscrollen durch einen zu lesenden Text. Diese Instrumente stellen keine effektive inhaltliche Kontrolle sicher, da sie leicht umgangen werden können. Kombiniert man diese Maßnahmen allerdings mit weiteren Elementen, wie beispielsweise einem entsprechend deutlichen Hinweisbanner (dazu sogleich) oder einer entsprechenden Zeitkomponente, in der bestimmte zu überprüfende Originalquellen betrachtet werden müssen, können sie allerdings eine effektive Kontrolle fördern.[349]
Nicht zu empfehlen ist demgegenüber jedoch die Verwendung von bewusst eingebauten Fehlern, die bei Nichtbeachtung Richterinnen und Richtern aufgezeigt werden und die Aufmerksamkeit erhöhen sollen. Zwar wären diese durchaus denkbar und zulässig. Sie sind allerdings deshalb nicht zu empfehlen, da auch die Akzeptanz und der Zuspruch von Nutzerinnen und Nutzern zu informationstechnischen Systemen einen ganz wesentlichen Gelingensfaktor für einen erfolgreichen Einsatz solcher Systeme bereithält.[350] Im Rahmen von Workshops, die mit Richterinnen und Richtern während unseres hier behandelten Projekts immer wieder durchgeführt worden sind, hat sich gezeigt, dass eine derartige Gestaltung von Kontrollmechanismen nicht auf großen Zuspruch und damit nicht auf Akzeptanz treffen dürfte.
Einen ähnlichen Gestaltungsansatz verfolgt auch Art. 14 Abs. 3 KI-VO, der Anbietern vorschreibt, durch Vorkehrungen Maßnahmen für die menschliche Aufsicht zu gewährleisten. Diese Aufsichtsmaßnahmen müssen den Risiken, dem Grad der

[349] Siehe auch *Dötterl*, RDi 2026, 414, 419, der mitunter auch ggf. ein erzwungenes Auseinandersetzen mit den Akten, Normen, einschlägiger Rechtsprechung und Literatur vorschlägt.
[350] Siehe hierzu ausführlich *Brodowski/Grabmair/Langer/Paschke/Rachut/Reuß/Treffers/Welpe/Wörner*, RDi 2026, 585ff.; siehe hierzu die Darstellung des *TITAN (Technologische Intelligenz zur Transformation, Automatisierung und Nutzerorientierung des Justizsystems)-Forschungsverbundes (Prof. Grabmair/TU Braunschweig)*, Darstellung Akzeptabilität, abrufbar unter https://www.tu-braunschweig.de/recht/digitl-hub/projekte/titan/vision/zielvorstellung (Abruf am 11.08.2026, 21:34 Uhr).

Autonomie und dem Kontext der Nutzung des Hochrisiko-KI-Systems angemessen sein, wobei sich diese Variablen typischerweise gegenseitig verstärken, dergestalt, dass etwa ein höheres Maß an Autonomie typischerweise höhere Risiken für die Schutzgüter von Gesundheit, Sicherheit oder Grundrechten (Art. 14 Abs. 2 KI-VO) nach sich zieht.[351] Die Vorkehrungen können anbieterseitig bereits in das Hochrisiko-KI-System vor dem Inverkehrbringen bzw. der Inbetriebnahme eingebaut werden, Art. 14 Abs. 3 lit. a KI-VO, oder der Anbieter kann Vorkehrungen bestimmen, die vor dem Inverkehrbringen oder der Inbetriebnahme geeignet sind, betreiberseitig umgesetzt zu werden, Art. 14 Abs. 3 lit. b KI-VO. Die grundsätzliche Wahlfreiheit der Alternativen-Wahl ist hierbei im Ergebnis danach auszurichten, dass die Vorkehrungen zu ergreifen sind, die das jeweilige Risiko am wirkungsvollsten zu senken oder auszuschließen geeignet sind.[352]
Abschließend sei auch auf den Vorschlag von *Gsell/Pfündl* hingewiesen, hinsichtlich eines verpflichtenden Rechtsgesprächs zwischen den Parteien und dem Gericht, um die Kontrolle der Ausgaben des informationstechnischen Systems durch den Richter bzw. die Richterin sicherzustellen.[353] Das Gericht werde durch das grundsätzlich verpflichtende Rechtsgespräch gezwungen, sich der Grundlagen und der Tragfähigkeit seiner eigenen, durch den Output informationstechnischer Systeme beeinflussten, vorläufigen Rechtsauffassung zu vergewissern.[354] Das verpflichtende Rechtsgespräch soll nach diesem Vorschlag nur mit der Zustimmung beider Parteien entbehrlich sein.[355]

(3) Sicherstellung von Kontrolle und Aufsicht im Hinblick auf psychologische Effekte

Insbesondere mit Blick auf die Gefahr der psychologischen Effekte empfiehlt sich ferner die **Aufklärung und Schulung der Richter und Richterinnen** sowie eine entsprechend sensibilisierte Gestaltung des informationstechnischen Systems.[356] Für Hochrisiko-KI-Systeme ergibt sich das Erfordernis der Ausbildung und Aufklärung explizit aus Art. 14 Abs. 4 KI-VO und Art. 26 Abs. 2 KI-VO.
Hinsichtlich der Aufklärung und Schulung könnte eine **periodische, verpflichtende Fortbildungsvorgabe** implementiert werden. Derartiges ist u. a. in der Fachanwaltschaft, vgl. § 15 Fachanwaltsordnung[357], bekannte Praxis und erscheint auch vor dem Hintergrund der Dynamik der Entwicklung im Bereich der künstlichen Intelligenz angezeigt; ebenso existieren bereits verpflichtende

[351] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 14 Rn. 85.
[352] BeckOK KI-Recht/*Buchner*, 7. Ed. 1.8.2026, KI-VO Art. 14 Rn. 40; Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 14 Rn. 86.
[353] *Gsell/Pfündl*, RDi 2026, 457, 463f., zwar ließe sich die Pflicht zur materiellen Prozessleitung gem. § 139 Abs. 1 S. 1 ZPO bereits dahingehend verstehen, eine Klarstellung im Gesetz sei dennoch geboten.
[354] *Gsell/Pfündl*, RDi 2026, 457, 464.
[355] *Gsell/Pfündl*, RDi 2026, 457, 464.
[356] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 119f., 242ff., 265ff.; *Schneider/Wundenberg/Sharaf*, ZHR 189 (2025), 259, 276; Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 13, 101f., 124.
[357] *DeutscheAnwaltAkademie Gesellschaft für Aus- und Fortbildung sowie Serviceleistungen mbH*, Fachanwaltsordnung, abrufbar unter https://www.anwaltakademie.de/fachanwaltslehrgaenge/fachanwaltsordnung (Abruf am 04.05.2026).

Fortbildungsveranstaltungen für Richterinnen und Richter.[358] Diese können um entsprechende Schulungen mit Inhalten zu verhaltenspsychologisch bedingten Verzerrungen[359], der Bias-Problematik beim Menschen, ergänzt werden, die die neuesten Erkenntnisse der Wissenschaft beinhalten. Insbesondere dem *Automation Bias* kann durch Schulungen begegnet werden.[360] Der Entwicklung effektiver Schulungskonzepte sollte wesentliches Augenmerk beigemessen werden.[361] Als weitere Maßnahmen im Rahmen des Wissensmanagements und -transfers können **Intervisionen** sowie der Austausch mit der Anwaltschaft (und anderen Akteuren des Justizsystems) genannt werden.[362]
Es ist ferner eine Vielzahl an Gestaltungsmöglichkeiten denkbar, die Wahrnehmung der Aufsicht und Kontrolle durch menschliche Richterinnen und Richter (systemseitig) zu unterstützen. Im Hochrisiko-KI-Bereich sind Anbieter sogar verpflichtet, technische Schutzmechanismen im KI-System zu verankern, um *Automation Bias* entgegenzuwirken.[363] Psychologischen Verzerrungen und Effekten kann u. a. damit entgegengewirkt werden, dass im Rahmen der Gestaltung der Nutzeroberfläche des Systems die Assistenzfunktion des Systems verdeutlicht wird und es den handelnden Akteur auf mögliche Fehlerquellen der Datengrundlage sowie der Verlässlichkeit der Ergebnisse aufmerksam macht.[364] Als Ausgestaltungsmöglichkeiten sind u. a. **Hinweisbanner mit der Aufforderung der kritischen Überprüfung** der Ergebnisse zu nennen.[365] Auch sollte auf die Positionierung von Anzeigen des Systems geachtet werden, da eine hervorgehobene Darstellung z. B. den *Automation Bias* erhöhend beeinflussen kann.[366] Neben den weiteren systemseitig vorzusehenden Instrumenten zur Sicherstellung von Aufsicht und Kontrolle kann ferner die Auferlegung von **Begründungslasten** (wobei an dieser Stelle mit der Begrifflichkeit der Begründung die Rechtfertigung der Übernahme von KI-Inhalten gemeint ist und die Begründungslast als Darlegung der inhaltlichen Überprüfung der übernommenen KI-Inhalte, insbesondere auf Vollständigkeit und Richtigkeit, z. B. durch Überprüfung der Ausgaben in den Quellen, zu verstehen ist; durch eine hinreichende Begründung kann die Kontrollpflicht ggf. als erfüllt angesehen werden) zur Gewährleistung einer hinreichend fundierten Auseinandersetzung mit den systemseitig ausgegebenen Ergebnissen insbesondere in sensiblen Bereichen angedacht werden.[367] Die

[358] Siehe hierzu bereits zuvor, Abschnitt C.II.2.a.bb(3).
[359] *Schneider/Wundenberg/Sharaf*, ZHR 189 (2025), 259, 276.
[360] *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 426.
[361] Vgl. *Schneider/Wundenberg/Sharaf*, ZHR 189 (2025), 259, 276.
[362] Vgl. die *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, Münchener Thesen zum Zivilprozess der Zukunft, abrufbar unter https://www.olg-duesseldorf.nrw.de/infos/Zivilprozess-der-Zukunft/Muenchener-Thesen-zum-Zivilprozess-der-Zukunft.pdf (Abruf am 26.02.2026, 18:27 Uhr), S. 9f., These B.4, zur Sicherung der qualitativen Bearbeitung von zivilprozessualen Verfahren.
[363] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 14 Rn. 106.
[364] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 14 Rn. 111; obgleich sie in zeitkritischen Situationen nur begrenzte Wirksamkeit aufweisen, Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 14 Rn. 112 mit Hinweis auf *Guijarro Santos*, ZfDR 2023, 23, 28ff.; *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 425f.
[365] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 296; Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 14 Rn. 112, wobei zu beachten ist, dass sie in zeitkritischen Situationen nur begrenzte Wirksamkeit aufweisen, a. a. O. mit Hinweis auf *Guijarro Santos*, ZfDR 2023, 23, 28ff.
[366] *Goddard/Roudsari/Wyatt*, Journal of the American Medical Informatics Association 19 (2012), 121, 125; hingegen können Übungen an Systemen mit absichtlich eingebauten Fehlern bzw. Fehlfunktionen den *Automation Bias* lindernd wirken, *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 426.
[367] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 296; Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 14 Rn. 113; *Nink*, Justiz und Algorithmen, 1. Aufl.

Begründung kann insbesondere darauf gerichtet sein, ob Anhaltspunkte für systemseitig unberücksichtigte Sachverhaltsinformationen im jeweils vorliegenden Fall ersichtlich sind.[368] Hinsichtlich des Erfordernisses der Begründung der Übernahme von Ergebnissen kann eine abgestufte Vorgehensweise verfolgt werden.[369] Hierbei bietet erneut die Funktionsmatrix[370], die eingehend im Rahmen des ersten Arbeitspakets (Milestone 1) des Forschungsprojekts „Künstliche Intelligenz und richterliche Entscheidungsfindung“ herausgearbeitet wurde, den Wegweiser für die Festsetzung des korrekten Maßes an Begründungslast, die dem Richter bzw. der Richterin auferlegt werden könnte. Bei rein administrativen Unterstützungsfunktionen, wie der Metadatenanalyse, könnte z. B. eine äußerst niedrigschwellige Begründungsanforderung bzw. sogar gar kein Begründungserfordernis ausreichen. Bei einem System, das automatisiert Akteninhalt in bereits vorformulierte Textbausteine einfügt, könnte aufgrund des mittleren Automatisierungsgrads auch ein nur mittleres Maß an Begründungserfordernis verlangt werden. Bei einer automatisierten und generativen Erstellung eigener Textvorschläge könnten – der Verortung der Funktion am äußeren Rand der Matrix folgend – höchste Begründungserfordernisse angelegt werden, wie eine inhaltliche Stellungnahme zur Übernahme des Textvorschlags, insbesondere die Vollständigkeit und Richtigkeit des Inhalts in hinreichend detailliertem Maß abdeckend. Hierbei ist jedoch auch zu betrachten, dass ein Begründungserfordernis bei der Übernahme von Ergebnissen des informationstechnischen Systems mit Aufwand verbunden ist, der Nutzerinnen und Nutzer ggf. davon abhalten kann, das System zu nutzen bzw. die Akzeptanz der Nutzerinnen und Nutzer verringert, da der Rationalisierungseffekt, der durch das Unterstützungssystem hervorgerufen werden sollte – gefühlt oder tatsächlich – neutralisiert oder ins Gegenteil verkehrt wird. Vor diesem Hintergrund sollte der Einsatz derartiger Begründungslasten vor der Implementierung hinreichend getestet werden, um ihren konkreten Einsatz insbesondere anhand der Faktoren der Rationalisierung durch die Unterstützung und Akzeptanz durch die Nutzerinnen und Nutzer zu untersuchen. Des Weiteren sollte die **Aufnahme von (einer Vielzahl an) Entscheidungsvarianten** (insbesondere in Funktionalitäten, die Textbausteine verwenden) in Betracht gezogen werden; diese sollten gleichberechtigt dargestellt werden, um nicht Default-Effekte auszulösen.[371] Insgesamt bietet es sich an, die (z. B. menschlichen, technischen sowie kontextuellen) Faktoren, die Bias begünstigen oder abmildern, bei der Implementierung entscheidungsunterstützender informationstechnischer Systeme zu beachten sowie deren Wechselwirkungen untereinander, um die Bias-Thematik adäquat adressieren zu können.[372] Hierzu sollten

2021, S. 425; *Martini/Botta/Nink/Kolain*, Automatisch erlaubt? Fünf Anwendungsfälle algorithmischer Systeme auf dem juristischen Prüfstand, 2020, S. 52.

[368] *Martini/Botta/Nink/Kolain*, Automatisch erlaubt? Fünf Anwendungsfälle algorithmischer Systeme auf dem juristischen Prüfstand, 2020, S. 52.

[369] *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 425.

[370] Bestimmung anhand erstens des **funktionalen Charakters der Tätigkeit** (rein administrative oder inhaltliche Funktionalität), zweitens des **Grades der Automatisierung,** sowie drittens der **Komplexität des jeweiligen Anwendungsfeldes.** Siehe hierzu ausführlich *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 12.

[371] *Rollberg*, Algorithmen in der Justiz, 1. Aufl. 2020, S. 49, 116 mit Hinweis auf *Berlit*, JurPC Web-Dok. 77/2012, Abs. 50; vgl. auch *Dötterl*, RDi 2026, 414, 419.

[372] Ausführlich zur Thematik des *Automation Bias* m. w. H. *Langer/Ruschemeier/Kolbe/Laux*, „Die KI wird’s schon richten“ – Eine psychologische Perspektive auf Automation Bias und richterliche Unabhängigkeit, in: Reuß/Laß (Hrsg.), Göttinger Kolloquien zur Digitalisierung des Zivilverfahrensrechts Kolloquien im Sommersemester 2024 | Kolloquien im Wintersemester 2024/2025, Göttingen 2026, S. 79ff.

die Erkenntnisse der psychologischen Forschung aktiv fortlaufend begleitet und deren Erkenntnisse implementiert werden.[373]

(4) Information über Funktionsweise und Limitierungen des Systems

Ein weiteres wesentliches Element der Sicherstellung von richterlicher Aufsicht und Kontrolle ist die Information der Systemnutzerinnen und Systemnutzer über die Funktionsweise und die Limitierungen des Systems.[374] Nur wenn die Arbeitsweise und die Limitierungen bekannt sind, kann auch die Kontrollnotwendigkeit von der Nutzerin bzw. von dem Nutzer konkret eingeschätzt und bewusst wahrgenommen werden (vgl. hierzu auch Art. 14 Abs. 4 KI-VO).
Bei den Fragen der Schaffung von Transparenz und Erklärung der (Ergebnisfindung der) KI ist in diesem Zusammenhang zu beachten, dass verschiedene Arten an Adressaten existieren, die eine unterschiedliche Erklärungstiefe benötigen sowie des Weiteren verschiedene Stufen der Erklärbarkeit existieren (z. B. erklärbare KI-Modelle, Post-Entscheidungs-Erklärungen, Datenerklärbarkeit).[375] Dies spiegelt sich auch in Art. 14 Abs. 4 lit. a KI-VO wider, nach dem Hochrisiko-KI-Systeme den Betreibern so zur Verfügung zu stellen sind, dass das Aufsichtspersonal die einschlägigen Fähigkeiten sowie Grenzen des Systems angemessen verstehen sowie den Betrieb ordnungsgemäß überwachen kann.[376]
Richterinnen und Richter, die vorwiegend mit informationstechnischen Unterstützungssystemen lediglich in der Praxis arbeiten werden, benötigen als zumeist **Nicht-KI-Experten** und **Domänenfachleute** (auf dem Gebiet des Rechts) eher eine **funktionale Erläuterung des Systems,** die dessen abstrakte Funktionsweise anschaulich und in für Nicht-Experten verständlicher Sprache beschreibt, den Zusammenhang zwischen Eingabe- und Ausgabedaten darstellt und Risiken für Fehlausgaben adressiert.[377] Dieses Bedürfnis der Information der Richterinnen und Richter sollte über verschiedene Kanäle adressiert werden. Hierbei versprechen **Schulungen und weitere KI-Kompetenzformate,** die Richterinnen und Richtern Schlüsselkompetenzen zu KI und KI-Systemen vermitteln, einen wesentlichen Baustein zu liefern.[378] Insbesondere Lehrgänge mit Zertifikatsabschlüssen sollten hierbei in Betracht gezogen werden (siehe hinsichtlich des Aspekts der Schulung von Richterinnen und Richtern bereits ausführlich zuvor insbesondere im Abschnitt

[373] Siehe insbesondere etwa zur Thematik des *Automation Bias* und der richterlichen Unabhängigkeit nach Art. 97 GG *Langer/Ruschemeier/Kolbe/Laux*, „Die KI wird’s schon richten“ – Eine psychologische Perspektive auf Automation Bias und richterliche Unabhängigkeit, in: Reuß/Laß (Hrsg.), Göttinger Kolloquien zur Digitalisierung des Zivilverfahrensrechts Kolloquien im Sommersemester 2024 | Kolloquien im Wintersemester 2024/2025, Göttingen 2026, S. 79ff.
[374] Vgl. *Köhler*, Richterliche Verantwortung und Künstliche Intelligenz, 1. Aufl. 2026, S. 137f.
[375] Vgl. Kaulartz/Braegelmann AI und Machine Learning-HdB/*Körner*, 1. Aufl. 2020, Kap. 2.4 Rn. 19ff.
[376] Im Kommissionsentwurf wurde noch das vollständige Verständnis des KI-Systems gefordert. Dies wurde in der Literatur als unrealistisch kritisiert. BeckOK KI-Recht/*Buchner*, 7. Ed. 1.8.2026, KI-VO Art. 14 Rn. 49f. m. w. N.
[377] Vgl. Kaulartz/Braegelmann AI und Machine Learning-HdB/*Körner*, 1. Aufl. 2020, Kap. 2.4 Rn. 20ff.
[378] Siehe hierzu BeckOK KI-Recht/*Kaufmann*, 7. Ed. 25.7.2026, KI-VO Art. 4 Rn. 35ff.; vgl. auch Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 124.

C.II.2.a.bb(3)).[379] Der Zugang zu den entsprechenden Informationen sollte jedoch zusätzlich auch intuitiv im Workflow eröffnet werden, wodurch Unsicherheiten im Arbeitsalltag sowie der Zeitverlust im Zusammenhang mit der Recherche der benötigten Informationen minimiert wird. Idealerweise sind derartige Informationen daher bereits **im System abrufbar**, etwa niederschwellig erreichbar unter einem Menüpunkt („Wie funktioniere ich und wo habe ich Limitierungen?!“) bzw. verlinkt im Hinweisbanner, das auf Limitationen und psychologische Effekte hinweist. Auch im Intranet bzw. auf der Webseite der Justiz können diese Informationen abgelegt werden, was letztlich auch dem Informationsinteresse der Öffentlichkeit und der betroffenen Beteiligten und Prozessvertreter entspricht. Im Rahmen des Hochrisiko-KI-Bereichs normiert thematisch überschneidend Art. 14 Abs. 4 lit. c KI-VO die anbieterseitige Pflicht, das System den Betreibern so zur Verfügung zu stellen, dass das Aufsichtspersonal adäquat in der Lage ist, die Ausgabe des Systems richtig zu interpretieren, i. S. eines Vermögens, die Informationen oder Empfehlungen des Systems richtig einordnen zu können.[380] Der Interpretationsrahmen umfasst hierbei neben den eigenen Kenntnissen über KI maßgeblich Informationen über die Funktionsweise des vorliegenden Systems sowie dessen angestrebten Einsatz.[381] Nach Art. 14 Abs. 4 lit. c KI-VO sind beispielsweise die vorhandenen Interpretationsinstrumente und -methoden bei der Interpretation zu berücksichtigen, was insbesondere werkseitig bereitgestellte Werkzeuge des Systems betrifft, wie z. B. Informationstexte oder auch graphische Aufbereitungen der für die Entscheidung relevanten Parameter.[382] Im Hochrisiko-KI-Bereich kann es mit Blick auf Art. 26 Abs. 2 KI-VO zudem sogar ggf. notwendig sein, dass die Systeme besondere Mechanismen aufweisen, die den beaufsichtigenden Menschen informieren bzw. beraten (siehe auch ErwGr. 73 S. 4 der KI-VO).[383]
Einen deutlich höheren Grad an Erklärungstiefe einschließlich mehrstufiger Erklärbarkeit, mithin **Experteninformationen**, benötigen hingegen KI-Spezialisten bzw. -Fachleute. Derartige Informationen sind daher insbesondere nutzbar und wertvoll für **Entwicklungs-** oder **Überwachungsgruppen**, die die Systementwicklung und dessen Betrieb begleiten, und sollten dort auch vorgesehen bzw. bereitgestellt werden. Für die Bereitstellung dieser Informationen bieten sich **Methoden der „Explainable AI“** an. Es besteht eine Vielzahl unterschiedlicher Ansätze in diesem Bereich, die Transparenz von KI-Systemen zu erhöhen, wobei u. a. modellunabhängige Ansätze, intrinsisch interpretierbare Ansätze sowie Post-hoc-Erklärungsmethoden identifiziert werden können.[384] Auch in diesem Bereich werden insbesondere Hybrid-Modelle aus regelbasierten Systemen und Modellen maschinellen Lernens vorgeschlagen sowie nutzerzentrierte KI-System-Designs mit

[379] Vgl. BeckOK KI-Recht/*Kaufmann*, 7. Ed. 25.7.2026, KI-VO Art. 4 Rn. 35; An dieser Stelle sei auf den Zertifikatsstudiengang: KI und Legal Tech an der Georg-August-Universität Göttingen hingewiesen, siehe hierzu Georg-August-Universität Göttingen, Neuer Zertifikatsstudiengang: KI und Legal Tech, abrufbar unter https://uni-goettingen.de/de/neuer+zertifikatsstudiengang%3A+ki+und+legal+tech/698161.html (Abruf am 24.10.2026, 17:04 Uhr).
[380] BeckOK KI-Recht/*Buchner*, 7. Ed. 1.8.2026, KI-VO Art. 14 Rn. 56.
[381] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 14 Rn. 119.
[382] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 14 Rn. 121ff.
[383] Martini/Wendehorst/*I. Eisenberger/Th. Binder*, 2. Aufl. 2026, KI-VO Art. 26 Rn. 25.
[384] *Pillai*, Enhancing transparency and understanding in AI decision-making processes. Iconic Research and Engineering Journals, 8 (2024) Nr. 1, S. 168ff.

intuitiver Bedienungsmöglichkeit und auf Interessengruppen zugeschnittenen Erläuterungen.[385]
Als „Explainable AI“-Ansätze können unter anderem sogenannte „saliency maps“ eingesetzt werden, die aufzeigen, welche Daten für das ausgegebene Ergebnis wichtig oder weniger ausschlaggebend waren.[386] Auch kann bei mittels Benutzereingaben gesteuerten informationstechnischen Systemen maschinellen Lernens die Nachvollziehbarkeit der Ausgaben durch eine Veranschaulichung der sog. „Aufmerksamkeitsmechanismen“ – diese werden genutzt, um relevante Informationen in den zugrundeliegenden Daten aufzufinden, zu gewichten und damit im Ergebnis Datenzusammenhänge auch über große Distanzen hinweg zu erfassen – erhöht werden, durch visuelle Darstellung der Aufmerksamkeitsverteilung auf unterschiedliche Eingabeteile.[387] Zudem ermöglichen u. a. sog. „counterfactual explanations“ die Untersuchung des Einflusses des Löschens oder der Veränderung von Eingabeteilen bei derart gesteuerten Systemen auf die Systemausgabe.[388] Verfügbare Techniken bieten jedenfalls bei komplexen Modellen in diesem Kontext derzeit wohl noch keine vollständigen Lösungen; die Verfolgung der Entwicklung dieser Methoden sowie entsprechende transparenzerhöhende Maßnahmen mit Blick auf die Verbindung von Input und Output sollten jedoch verfolgt werden.[389] Solche Ansätze bestehen derzeit insbesondere für bestimmte Modelle;[390] idealiter sollte die Entwicklung dieser Ansätze zur Transparenzerhöhung parallel zur Entwicklung der spezifischen informationstechnischen Systeme für die richterliche Entscheidungsfindung mit aufgebaut werden.
Um die Erklärbarkeit von KI-Modellen zu erhöhen, kann zudem eine selbsterklärende Ausgestaltung der KI-Modelle angestrebt werden, die neben dem Ausgabeergebnis auch darauf trainiert wurden, eine Begründung zu diesem zu liefern.[391] Hinsichtlich Post-Entscheidungs-Erklärungsmethoden besteht eine Vielzahl an Ansätzen; im Wesentlichen lassen diese sich aufteilen in: das Ziel, das über das Modell erklärt werden soll; die Gründe, die zu einem Erklärungsbedarf führen; die Erklärungsart (insbesondere Text, Beispiele, Deskription, Bilder, Zahlen- und Konfidenzwerte); der Prozess, der die Erklärung im Nachgang generiert und zur Verfügung stellt.[392] Eine möglichst umfassende Betrachtung dieser Ansätze bzw. deren Implementierung führt dabei zu einer steigenden Transparenz.

cc. Zusammenfassung

Beim entscheidungsunterstützenden Einsatz können informationstechnische Systeme derzeit ebenfalls nur in hinreichend einfach gelagerten Fällen die Einhaltung der rechtlichen Anforderungen sicherstellen. Der einsetzende Richter oder die einsetzende Richterin muss die Ergebnisse des Systems kontrollieren und bei Bedarf korrigieren. **Richterinnen und Richter tragen die Letztverantwortung für die Gesetzeskonformität der Entscheidung** (sowie insbesondere dafür, dass alle

[385] *Pillai*, Enhancing transparency and understanding in AI decision-making processes. Iconic Research and Engineering Journals, 8 (2024) Nr. 1, S. 168ff.
[386] Kaulartz/Braegelmann AI und Machine Learning-HdB/*Rühl*, 1. Aufl. 2020, Kap. 14.1 Rn. 20.
[387] *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 178; *Hacker/Cordes/Berz*, GRUR 2024, 1777, 1777, 1782.
[388] *Hacker/Cordes/Berz*, GRUR 2024, 1777, 1782.
[389] *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 179; *Hacker/Cordes/Berz*, GRUR 2024, 1777, 1782.
[390] *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 179.
[391] Kaulartz/Braegelmann AI und Machine Learning-HdB/*Körner*, 1. Aufl. 2020, Kap. 2.4 Rn. 24.
[392] Kaulartz/Braegelmann AI und Machine Learning-HdB/*Körner*, 1. Aufl. 2020, Kap. 2.4 Rn. 25.

verfahrensrelevanten Informationen zutreffend erfasst und verarbeitet wurden und sich ein systemseitiger Bias und systemseitige Halluzinationen nicht ergebnisrelevant niederschlagen). Hierfür sind **effektive Kontroll- und Korrekturmöglichkeiten einzuräumen**. Die notwendige Kontrollintensität steigt mit der Komplexität des Anwendungsfelds, dem Grad der Automatisierung und der inhaltlichen Funktionalität des Systems. Psychologischen Effekten wie einem „Automation Bias“ ist vorzubeugen. Die Richterinnen und Richter sind über die Funktionsweise und Limitierungen des Systems zu informieren.

Zur **Sicherstellung substantieller Kontroll- und Korrekturmöglichkeiten** bestehen mehrere Möglichkeiten. Allgemein sollte Richterinnen und Richtern eine möglichst weitgehende Nachvollziehbarkeit ermöglicht werden hinsichtlich der Kriterien und Aspekte für die Ergebnisgenerierung. Empfohlen wird eine herkunftssensitive Benutzeroberfläche, die Informationen möglichst im originalen Wortlaut wiedergibt und die originalen Quellen(passagen) per Klick sowie ohne Medienbruch sichtbar und abrufbar macht und eine niederschwellige Korrekturmöglichkeit bei Bedarf bietet. Sinngemäße Zusammenfassungen sind – etwa bei umfangreichen Schriftsätzen – als Systemausgabe nicht gänzlich abzulehnen, bringen aber weitergehende Kontrollerfordernisse mit sich. Idealerweise ist das System in die e-Akte integriert. Von Scores, welche die Sicherheit einer Ausgabe angeben, ist abzuraten. Diese sind ihrerseits nicht verlässliche Systemausgaben und können durch eine fälschlich hohe Sicherheitsangabe zum Verzicht auf eine tatsächlich gebotene Prüfung verleiten. Halluzinationen schließen den entscheidungsunterstützenden Einsatz informationstechnischer Systeme (anders als den entscheidungsersetzenden) wegen des richterlichen Kontrollerfordernisses und der richterlichen Letztverantwortung nicht aus, sollten aber etwa durch RAG-Systeme und Konsistenzkontrollen vermieden werden, um Rationalisierungseffekte nicht zu beeinträchtigen. Befürwortet wird zudem eine strukturierte Kommunikation samt einer möglichen Begrenzung des Vortragsumfangs, die die Genauigkeit der Systemergebnisse erhöht sowie die Aufsicht und Kontrolle erleichtert.

Für die Wahrnehmung der bereits aus der Gesetzesbindung folgenden Kontrollpflicht ist ein Kontrollmodell zu empfehlen, das abhängig von der Komplexität des Anwendungsfeldes, dem Grad der Automatisierung und der Funktionalität des Systems gestaltet ist. Bei Funktionen im mittleren bis höchsten Bereich des Kontrollerfordernisses kommt etwa eine **verpflichtende Stichprobenprüfung** der Systemergebnisse vor der Signatur in Betracht. Zwar ersetzt dies nicht die nach Art. 20 Abs. 3 GG gebotene vollständige Prüfung der Systemergebnisse, setzt einem Technikübervertrauen aber einen Gegenimpuls entgegen. Die Kombination mehrerer Kontrollmechanismen kann für eine effektive inhaltliche Kontrolle vorteilhaft sein. Aus Akzeptanzgründen empfiehlt es sich nicht, bewusste Fehler einzubauen.

Hinsichtlich der Sicherstellung von Kontrolle und Aufsicht im Hinblick auf psychologische Effekte empfehlen sich die **Aufklärung und Schulung**, ggf. einschließlich periodischer, verpflichtender Fortbildungsvorgaben, eine sensibilisierte Benutzeroberflächengestaltung, z. B. mit einem Hinweisbanner mit der Aufforderung zur kritischen Überprüfung, eine bedachte Positionierung von **Systemanzeigen** und die gleichberechtigte **Darstellung** mehrerer Entscheidungsvarianten. Weiterhin sind nach der Funktionsmatrix **abgestufte Begründungslasten** für die Übernahme von KI-Inhalten in diesem Zusammenhang eine mögliche Vorgehensweise.

Richterinnen und Richter sind adressatengerecht über die Funktionsweise und Limitierungen des Systems zu informieren. Sie benötigen vor allem eine funktionale, verständliche Erläuterung. Schulungen und weitere **KI-Kompetenzformate** sowie die unmittelbare Abrufbarkeit derartiger Informationen im System erscheinen hierbei

vorteilhaft. Entwicklungs- und Überwachungsgruppen benötigen hingegen einen deutlich größeren Grad an Erklärungstiefe. Methoden der „Explainable AI“ bieten hierfür Ansätze, bei komplexen Modellen derzeit aber wohl noch keine vollständigen Lösungen. Ihre Entwicklung sollte verfolgt und möglichst parallel zu den Systemen für die richterliche Entscheidungsfindung aufgebaut werden.

3. Schutz informationstechnischer Systeme vor Störungen

Weitere zentrale rote Linien betreffen den Schutz informationstechnischer Systeme vor Störungen, denn diese können Einfluss auf das Ausgabeergebnis dieser Systeme nehmen und zeitigen damit unmittelbar Auswirkungen auf die Aufgabenerledigung.

a. Übergreifende rote Linien für jeden Einsatz

Die in dem ersten Arbeitspaket (Milestone 1) des Forschungsprojekts „Künstliche Intelligenz und richterliche Entscheidungsfindung“ herausgearbeiteten roten Linien in diesem Zusammenhang sind v. a. übergreifender Natur. Sie sind daher sowohl für den entscheidungsersetzenden als auch für den entscheidungsunterstützenden Einsatz relevant.[393] Hierzu zählen:
Es ist erforderlich, den Schutz des informationstechnischen Systems vor internen und externen Störungen sicherzustellen, um die ordnungsgemäße Funktion und den gesetzmäßigen Betrieb des Systems zu ermöglichen.[394] **Interne und externe Störungen** des informationstechnischen Systems können z. B. durch einen Defekt der verwendeten Software (z. B. Programmierfehler), einen Stromausfall oder einen Hackerangriff auftreten.[395] Es muss systemseitig oder durch sonstige Vorkehrungen sichergestellt sein, dass Dritte keinen unzulässigen unmittelbaren oder mittelbaren Einfluss auf die Entwicklung, Gestaltung und Kontrolle des informationstechnischen Systems nehmen können; die Entscheidungsfindung des informationstechnischen Systems darf sich letztlich nur an der Gesetzesbindung orientieren.[396] Dies ergibt sich insbesondere aus den Art. 92, 97, 101 Abs. 1 Satz 2 GG[397] und aus Art. 6 Abs. 1 EMRK[398] sowie aus Art. 47 Abs. 1, Abs. 2 GRCh[399].

b. Handlungsoptionen und -empfehlungen

aa. Schutz vor internen oder externen Störungen; Sicherstellung der ordnungsgemäßen Funktion und des gesetzmäßigen Betriebs

Die Sicherheit der informationstechnischen Systeme ist ein wesentlicher Aspekt ihres Einsatzes, sie ist nicht nur singulär, sondern vielmehr über den ganzen Lebenszyklus des informationstechnischen Systems zu betrachten.[400] Hierbei sind insbesondere **vier Hauptareale** zu beachten, die Sicherheit des Designs, die Sicherheit der

---

[393] Siehe hierzu ausführlich *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025.
[394] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 118, 131ff., 167ff., 187ff., 198ff., 216ff., 242ff., 265ff.
[395] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 118, 168, 198ff., 218ff., 245f., 265ff.
[396] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 131ff., 167ff., 187ff., 216ff., 245f., 265ff.
[397] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 167ff., 187ff., 216ff.
[398] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 245f.
[399] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 265ff.
[400] *National Cyber Security Centre et al.*, Guidelines for secure AI system development, abrufbar unter https://www.ncsc.gov.uk/sites/default/files/documents/Guidelines-for-secure-AI-system-development.pdf (Abruf am 23.04.2026, 13:14 Uhr), S. 5.

Entwicklung, die Sicherheit der Bereitstellung sowie der Komplex sicherer Betrieb und Wartung.[401]

Im Rahmen der **Sicherheit des Designs** sollten als Schlüsselfaktoren die Sensibilisierung der Mitarbeiter für Gefahren und Risiken, die Bewertung der Risiken für das System, eine sichere, funktionale und leistungsfähige Systemgestaltung (umfassend die Sicherheit der Lieferketten, sowie die Frage, ob eine interne Entwicklung oder der Einsatz externer Komponenten vorzugswürdig ist, Einschränkungen ggf. von KI ausgelösten Aktionen, die Modellierung von Nutzerinteraktionen unter Beachtung KI-spezifischer Risiken) und die – aufgrund des ständigen Fortschritts der wissenschaftlichen Erkenntnisse zur KI-Sicherheit und KI-Risiken regelmäßig neu zu evaluierende – Berücksichtigung von Sicherheitsvorteilen und -kompromissen bei der Auswahl des KI-Modells eine Rolle spielen.[402]

Im Rahmen der **Sicherheit der Entwicklung** sind insbesondere relevant die Sicherheit der Lieferkette (über den ganzen Lebenszyklus des informationstechnischen Systems hinweg, wobei bei nicht In-house-Lösungen Hardware und Software von hohem Sicherheitsstandard gewählt werden sollte), die Identifikation, Nachverfolgung sowie der Schutz der mit dem informationstechnischen System einhergehenden Vermögenswerte, eine umfassende – insbesondere Sicherheitsaspekte umfassende – Dokumentation der Daten, Modelle und Prompts sowie eine genaue Planung des Projekts und Implementierung der Best Practices.[403]

Bei der **Sicherheit der Bereitstellung** des Systems sollte auf eine sichere Infrastruktur für jeden Teil des systemischen Lebenszyklus mit Zugriffskontrollen sowie Abtrennungen kritischer Bereiche, einen kontinuierlichen Schutz des Systems anhand von Best Practices aus der Cybersicherheit, Schnittstellen-Abfrage-Kontrollen, Verschlüsselungen und – der Sensibilität des justitiellen Einsatzes entsprechend – hohen Datenschutzmaßnahmen, Wert gelegt werden.[404] Wozu bereits ungeplante Eingriffe (z. B. durch Bauarbeiten) in die digitale Infrastruktur der Justiz führen können, hat sich gerade erst jüngst in der niedersächsischen Justiz gezeigt, wo durch einen Bagger unfreiwillig eine Hauptdatenleitung zertrennt worden war und die Justiz damit fast nahezu vollständig offline und damit arbeitsunfähig war.[405] Derartige Situationen sind vor dem Hintergrund der Bedeutung der Justiz im demokratischen System von besonderer Schwierigkeit und sollten durch sorgfältige Planung und Absicherung (etwa durch Ersatzdatenleitungen) vermieden werden. Dies sollte weiter von iterativ aktualisierten und auf verschiedene Sicherheitsinzidenzen ausgerichteten Maßnahmenplänen für Sicherheitsfälle, Offline-Backups und umfassend geschultem

---

[401] *National Cyber Security Centre et al.*, Guidelines for secure AI system development, abrufbar unter https://www.ncsc.gov.uk/sites/default/files/documents/Guidelines-for-secure-AI-system-development.pdf (Abruf am 23.04.2026, 13:14 Uhr), S. 5.

[402] *National Cyber Security Centre et al.*, Guidelines for secure AI system development, abrufbar unter https://www.ncsc.gov.uk/sites/default/files/documents/Guidelines-for-secure-AI-system-development.pdf (Abruf am 23.04.2026, 13:14 Uhr), S. 9f. m. w. H.

[403] *National Cyber Security Centre et al.*, Guidelines for secure AI system development, abrufbar unter https://www.ncsc.gov.uk/sites/default/files/documents/Guidelines-for-secure-AI-system-development.pdf (Abruf am 23.04.2026, 13:14 Uhr), S. 12f. m. w. H.

[404] *National Cyber Security Centre et al.*, Guidelines for secure AI system development, abrufbar unter https://www.ncsc.gov.uk/sites/default/files/documents/Guidelines-for-secure-AI-system-development.pdf (Abruf am 23.04.2026, 13:14 Uhr), S. 14 m. w. H.

[405] Dazu der Bericht bei *Die Zeit*, Ein Bagger, ein Kabel – und die Justiz steht still, abrufbar unter https://www.zeit.de/news/2026-06/10/kabel-beschaedigt-landesweite-it-stoerung-trifft-gerichte (Abruf am 13.07.2026).

Personal flankiert werden.[406] Die Resilienz justizieller Systeme ist von überragendem Interesse.

Um effiziente sowie effektive Maßnahmen der **Informationssicherheit** maßzuschneidern, sollte eine Risikobewertung für das jeweils zum Einsatz intendierte informationstechnische System vorgenommen werden; es sollte sichergestellt sein, dass nur Befugte in die Konzeption, Programmierung, Nutzung, Überwachung und das Training eingebunden sind.[407] Durch ein entsprechend angepasstes Informationssicherheitsmanagementsystem sollten die Teilaspekte der Informationssicherheit im hinreichenden Umfang gewährleistet sowie aufrechterhalten werden.[408] Bei der Risikoanalyse sollten bestehende, einschlägige aktuelle ISO-Normen, etwa die ISO-Norm 27005, 31000, aber auch 23894 sowie aktuelle BSI-Standards wie der BSI-Standard 200-3 beachtet werden.[409] Informationssicherheit und -management sind keine einmalige Aufgabe, sondern ein dauerhafter Prozess mit dem Ziel, das angestrebte Sicherheitsniveau zu erreichen, aufrechtzuerhalten und zu verbessern, wozu eine regelmäßige Überprüfung der Maßnahmen sowie ihre Analyse auf Praxisnähe und korrekte Umsetzung erforderlich sind.[410]

Im Rahmen von Risikoanalysen kann freilich nicht jede erdenkliche Situation betrachtet werden, vielmehr stehen die wahrscheinlichsten Gefahren sowie plausibelsten Lösungen im Vordergrund, wobei in der Praxis die „80:20“-Regel angesetzt wird.[411] Mit Blick auf die oftmals sensiblen Daten[412], die in der Justiz Gegenstand sind, und massiven Auswirkungen einer Systemkorruption[413] sollten freilich höchste Sicherheitsraten und entsprechende Regeln angestrebt werden. Zentrale Maßnahmen sind klare Klassifizierungen der Sensibilität des jeweiligen Bereichs, klare Zuständigkeiten und Verantwortungszuweisungen, insbesondere auch mit Blick auf Roh- und Trainingsdaten, die mit Zugangsbeschränkungen, -kontrollen und (logischen sowie physischen) -steuerungen, in ihrer Intensität entsprechend der Risiken, und deren Überprüfung.[414] Auch Gefahren durch ausscheidendes Personal oder infolge auslaufender Verträge sind in den Blick zu nehmen, sowie ein Dokumentationssystem, insbesondere für signifikante Ereignisse, anzudenken und (kryptographische) Verschlüsselungsverfahren sowie mehrstufige Sicherheitsvorkehrungen sollten implementiert werden, damit bei Überwindung einer Sicherheitshürde nicht bereits eine Störung eintritt.[415] Auch bieten wiederum KI-Systeme im Rahmen von Risikoanalysen großes Potential, erhebliche Risikofaktorenmengen einschließlich Wechselwirkungen zu verarbeiten sowie bei der Suche nach optimalen Lösungen, auch auf gänzlich neuen Wegen.[416]

Sicherheitsfragen sowie **Angriffe auf informationstechnische Systeme** sind der Justiz nicht unbekannt.[417] Richter sind in Fragen der IT-Sicherheit grundsätzlich nicht

---

[406] *National Cyber Security Centre et al.*, Guidelines for secure AI system development, abrufbar unter https://www.ncsc.gov.uk/sites/default/files/documents/Guidelines-for-secure-AI-system-development.pdf (Abruf am 23.04.2026, 13:14 Uhr), S. 14f. m. w. H.
[407] Kaulartz/Braegelmann AI und Machine Learning-HdB/*Kipker/Müller*, 1. Aufl. 2020, Kap. 8.6 Rn. 15.
[408] Kaulartz/Braegelmann AI und Machine Learning-HdB/*Kipker/Müller*, 1. Aufl. 2020, Kap. 8.6 Rn. 15.
[409] Kaulartz/Braegelmann AI und Machine Learning-HdB/*Kipker/Müller*, 1. Aufl. 2020, Kap. 8.6 Rn. 15.
[410] Kaulartz/Braegelmann AI und Machine Learning-HdB/*Kipker/Müller*, 1. Aufl. 2020, Kap. 8.6 Rn. 24.
[411] Kaulartz/Braegelmann AI und Machine Learning-HdB/*Kipker/Müller*, 1. Aufl. 2020, Kap. 8.6 Rn. 16.
[412] *Vogelgesang*, jM 2018, 2, 3.
[413] Vgl. *Wolff*, Algorithmen als Richter, 1. Aufl. 2022, S. 170.
[414] Kaulartz/Braegelmann AI und Machine Learning-HdB/*Kipker/Müller*, 1. Aufl. 2020, Kap. 8.6 Rn. 17ff.
[415] Kaulartz/Braegelmann AI und Machine Learning-HdB/*Kipker/Müller*, 1. Aufl. 2020, Kap. 8.6 Rn. 18, 20, 22ff.
[416] *Wischmeyer*, AöR 143 (2018), 1, 16.
[417] *Vogelgesang*, jM 2018, 2, 2, 6f.

weisungsfrei.[418] Zudem sind Rechtsprechungsbeeinflussungsversuche (etwa Korruption oder Erpressung) bereits aus der analogen Welt bekannt.[419] Die informationstechnische Gefährdungslage ist gleichwohl bereits vom Sprung der analogen Aktenarbeit zur digitalen Aktenarbeit quantitativ sowie qualitativ erheblich angestiegen.[420] Mit Blick auf die mögliche Rechtsprechungsbeeinflussung durch manipulierte informationstechnische Systeme bei der richterlichen Entscheidungsfindung hat sich diese Lage noch einmal intensiviert.[421] Dies bedeutet jedoch nicht den Ausschluss informationstechnischer Systeme bei der richterlichen Entscheidungsfindung, sondern vielmehr die Pflicht, möglichst umfassende Maßnahmen zur Erreichung eines ausreichenden Sicherheitsniveaus zu ergreifen.[422] **Präventive digitale Absicherungen mit umfangreichen Testphasen und Evaluierungen vor dem Ausrollen der Systeme** sollten als wesentliche Fokuspunkte eingesetzt werden.[423] Hierbei kommen insbesondere **öffentliche Hacking-Tests** in Betracht, um die Belastbarkeit und Widerstandsfähigkeit gegen Angriffe sicherzustellen.[424] Auch sollten **nur authentifizierte Geräte auf die informationstechnischen Systeme zugreifen** dürfen und **keine Übersendungen von Daten über private Kanäle** von diesen zugelassen werden.[425] Neben digitalen Sicherheitsmaßnahmen sind natürlich auch noch „analoge“ Maßnahmen einzuhalten, etwa, dass der Zutritt Dritter zu Diensträumen beschränkt und kontrolliert wird, Bürotüren gewissenhaft auch bei nur kurzer Abwesenheit abgeschlossen werden, Rechner nicht in der Nähe von Besucherplätzen installiert werden und Rechner auch selbst bei Verlassen des Büros gesperrt werden.[426]
Mit dem Einsatz informationstechnischer Systeme sollte zum Schutz vor internen oder externen Störungen sowie zur Sicherstellung der ordnungsgemäßen Funktion und des gesetzmäßigen Betriebs ein **präventives, umfassendes Sicherheits- und Cyber-Resilienzkonzept** einhergehen, das die Vorbereitung, Prävention, Erkennung und Reaktion, sowie die zügige Wiederherstellungsmöglichkeit der Arbeitsfähigkeit und Anpassung an neue Bedrohungslagen miteinander verzahnt.[427] Teile dieses Konzepts sollten insbesondere die Kombination technischer, organisatorischer sowie personeller Faktoren und Maßnahmen sein sowie Maßnahmen gegen unberechtigten Zugriff, regelmäßige Updates und Patches des Systems, elaborierte Maßnahmenpläne, regelmäßige Backups und Datenwiederherstellungsmechanismen.[428]
Zwar können Systeme künstlicher Intelligenz ihrerseits zur Informationssicherheit beitragen[429], jedoch sind auch ihre Eigenheiten sowie speziellen Gefahren im Rahmen des Schutzes vor Störungen sowie für die Sicherstellung der ordnungsgemäßen Funktion und des gesetzmäßigen Betriebs zu beachten. So sind etwa **effektive**

[418] *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 369 Fn. 43.
[419] *Horn*, ZZP 136 (2023), 459, 480.
[420] *Vogelgesang*, jM 2018, 2, 4.
[421] Vgl. *Wolff*, Algorithmen als Richter, 1. Aufl. 2022, S. 170.
[422] *Horn*, ZZP 136 (2023), 459, 480.
[423] Vgl. *Wolff*, Algorithmen als Richter, 1. Aufl. 2022, S. 170.
[424] *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 369f. Fn. 43.
[425] Vgl. *Vogelgesang*, jM 2018, 2, 4ff., die die Geräteauthentifizierung in Bezug auf mobile Datenträger, die an Dienstrechner angeschlossen werden, ausführt.
[426] *Vogelgesang*, jM 2018, 2, 6.
[427] Vgl. hierzu *AlHidaifi/Asghar/Ansari*, A Survey on Cyber Resilience: Key Strategies, Research Challenges, and Future Directions, ACM Computing Surveys 56 (2024), Nr. 8, Art. 196, Art. 196ff.
[428] Vgl. hierzu *Beretas*, Information Systems Security, Detection and Recovery from Cyber Attacks, Universal Library of Engineering Technology 1 (2024), Nr. 1, S. 27.
[429] Kaulartz/Braegelmann AI und Machine Learning-HdB/Kipker/*Müller*, 1. Aufl. 2020, Kap. 8.6 Rn. 9ff.

Maßnahmen gegen sogenannte Prompt-Injections als wesentliche Sicherheitslücke bei generativer künstlicher Intelligenz, insbesondere bei LLMs zu beachten.[430] Prompt-Injections sind Angriffe, bei denen fremde Anweisungen gegeben werden, die eine KI dazu bringen sollen, von der richtigen Anweisung abzuweichen.[431] Diese können z. B., durch für das menschliche Auge versteckte Weise, in Schriftsätzen niedergelegt werden[432] und so das informationstechnische System bei der richterlichen Entscheidungsfindung beeinflussen mit Auswirkungen (auch) auf den/die menschlichen Richter/in bzw. die Entscheidungsfindung sowie -fällung. Je nach Konstellation können Prompt-Injections strafrechtliche sowie zivilrechtliche Sanktionen auslösen.[433] Dies setzt jedoch voraus, dass diese erst einmal bemerkt werden, was je nach Grad der Subtilität etwa bei der Manipulation entscheidungsunterstützender Systeme bei der richterlichen Entscheidungsfindung unter Umständen schwierig sein kann. Obgleich keine absolute Sicherheit bezüglich Prompt-Injections beim Einsatz entsprechender Systeme erreichbar scheint, sollten weitestmöglich Maßnahmen gegen diese ergriffen werden.[434] Die Nutzung strukturierter Eingabemasken, z. B. auf einer Justizplattform, macht beispielsweise solche Prompt-Injections sichtbar, da der eingegebene Text sichtbar ist. Ein Versehen z. B. eines hochzuladenden PDF-Dokuments mittels für den Menschen nicht lesbaren Texts bliebe aber möglich. Hier könnten beispielsweise Filteranwendungen helfen, die injizierten Text beispielsweise aufdecken und vor der Verarbeitung durch KI herausfiltern. Ein ergänzendes Instrument wäre die entsprechende Anpassung im Systemprompt dahingehend, dass bestimmte Textbestandteile (z. B. in weißer Farbe auf weißem Grund gestalteter Text) nicht verarbeitet werden sollen, was jedoch nicht vollständig zuverlässig gelingen wird, wenn das System hierauf nicht trainiert worden ist.
Auch ist zu beachten, dass LLMs selbst ein erhebliches Gefährdungspotential mit Blick auf informationstechnische Systeme bergen. Zunehmend werden KI-Modelle genutzt, um Hacks auszuführen; so nutzte Anfang des Jahres 2026 ein Hacker die Claude-KI von Anthropic, um in mexikanische Behördennetzwerke einzudringen.[435] Dies folgt einem Trend, in dem KI zur (staatlichen) Cyberspionage und -kriminalität eingesetzt wird.[436] Die Risiken für die Cybersicherheit durch immer leistungsstärkere LLMs sollten eng begleitet werden sowie entsprechende Maßnahmen getroffen werden; dieser Trend ist nicht zu unterschätzen, so hält das US-Unternehmen Anthropic das KI-Modell

[430] *Kuhlmann*, Prompt Injections – Betrügen für Fortgeschrittene, abrufbar unter https://www.lto.de/recht/hintergruende/h/prompt-injections-erklaerung-ist-das-strafbar (Abruf am 23.04.2026, 21:16 Uhr).
[431] Vgl. *OpenAI*, Prompt Injections verstehen: eine neuartige Sicherheitsherausforderung, abrufbar unter https://openai.com/de-DE/index/prompt-injections/ (Abruf am 23.04.2026, 21:19 Uhr).
[432] *Kuhlmann*, Prompt Injections – Betrügen für Fortgeschrittene, https://www.lto.de/recht/hintergruende/h/prompt-injections-erklaerung-ist-das-strafbar (Abruf am 23.04.2026, 21:16 Uhr).
[433] *Kuhlmann*, Prompt Injections – Betrügen für Fortgeschrittene, abrufbar unter https://www.lto.de/recht/hintergruende/h/prompt-injections-erklaerung-ist-das-strafbar (Abruf am 23.04.2026, 21:16 Uhr).
[434] Vgl. hierzu *Kosinski*, So verhindern Sie Prompt-Injection-Angriffe, abrufbar unter https://www.ibm.com/de-de/think/insights/prevent-prompt-injection (Abruf am 23.04.2026, 21:34 Uhr).
[435] *Knobloch*, Claude: KI-Chatbot für Cyberangriff auf mexikanische Regierung genutzt, abrufbar unter https://www.heise.de/news/Claude-KI-Chatbot-fuer-Cyberangriff-auf-mexikanische-Regierung-genutzt-11190396.html (Abruf am 23.03.2026, 21:45 Uhr).
[436] *Knobloch*, Claude: KI-Chatbot für Cyberangriff auf mexikanische Regierung genutzt, abrufbar unter https://www.heise.de/news/Claude-KI-Chatbot-fuer-Cyberangriff-auf-mexikanische-Regierung-genutzt-11190396.html (Abruf am 23.03.2026, 21:45 Uhr).

„Claude Mythos“ vor der allgemeinen Zugänglichkeit zurück.[437] Laut eigener Angabe habe das System bereits „tausende“ schwerwiegende Sicherheitsschwachstellen in jedem oft eingesetzten Betriebssystem sowie Webbrowser gefunden; die öffentliche Verfügbarkeit derart wirkmächtiger KI ergibt Fragen zur nationalen und europäischen Sicherheit und Souveränität laut Claudia Plattner, der Präsidentin des BSI.[438]

bb. Etablierung eines Risikomanagementsystems

Auch sollten die Aspekte der KI-Verordnung im Blick gehalten werden, insbesondere der Vorgaben zum **Risikomanagementsystem** nach Art. 9 KI-VO für Hochrisiko-Systeme, unabhängig von deren Vorliegen. Art. 9 Abs. 2 S. 1 KI-VO beschreibt das Risikomanagementsystem wie folgt: „*Das Risikomanagementsystem versteht sich als ein kontinuierlicher iterativer Prozess, der während des gesamten Lebenszyklus eines Hochrisiko-KI-Systems geplant und durchgeführt wird und eine regelmäßige systematische Überprüfung und Aktualisierung erfordert.*“ Die Implementierung eines solchen sollte stets erfolgen. Durch diese Implementierung unabhängig von der Hochrisiko-Einordnung besteht nicht die Gefahr einer Falscheinordnung und damit eines Verstoßes gegen die KI-VO. Mit Blick auf die Neuartigkeit der Regelungen der KI-VO und den damit einhergehenden noch (insbesondere höchstrichterlich) ungeklärten Rechtsfragen, insbesondere zur Abgrenzung, ob ein Hochrisiko-KI-System vorliegt oder nicht, erscheint dies vorteilhaft. Jedenfalls aber sollte zumindest eine Bund/Länder-übergreifende Verständigung auf Expertengrundlage zur Auslegung – insbesondere zur Frage der Einordnung von Systemen als Hochrisiko-KI-Systemen – erarbeitet und etabliert werden.[439] Die Vorgaben des Art. 9 KI-VO gliedern sich in die allgemeine Pflicht, ein Risikomanagementsystem einzurichten, anzuwenden, zu dokumentieren und aufrechtzuerhalten, gem. Art. 9 Abs. 1 KI-VO. Hiermit wird vorgegeben, dass derartige Abläufe und Strukturen eingerichtet werden müssen, die den Mindestanforderungen der Verordnung entsprechen, die konkrete Art und Weise der Umsetzung wird hingegen nicht vorgegeben.[440]
Das Risikomanagementsystem lässt sich in vier verschiedene, kontinuierlich zu wiederholende Schritte ordnen, die **Ermittlung und Analyse vernünftigerweise vorhersehbarer Risiken** (Art. 9 Abs. 2 S. 2 lit. a) KI-VO), **deren Abschätzung und Bewertung** (Art. 9 Abs. 2 S. 2 lit. b) KI-VO), die **Evaluation der Produktbeobachtung nach dem Inverkehrbringen des KI-Systems** (Art. 9 Abs. 2 S. 2 lit. c) KI-VO) sowie das **Treffen geeigneter und gezielter Risikomanagementmaßnahmen** (Art. 9 Abs. 2 S. 2 lit. d) KI-VO), deren Ausgestaltung sich aus Art. 9 Abs. 3–5 KI-VO ergibt.[441] Art. 9 Abs. 3 KI-VO konkretisiert die im Rahmen des Risikomanagements zu berücksichtigenden Risiken. Art. 9 Abs. 4 KI-VO beschreibt weitere Anforderungen mit Blick auf Hochrisiko-KI-Systeme, die seitens des Anbieters bei der Wahl der Risikomanagementmaßnahmen

[437] *ZDF*, „Claude Mythos“ von Anthropic: Neues KI-Modell: BSI zeigt sich besorgt, abrufbar unter https://www.zdfheute.de/politik/deutschland/ki-anthropic-claude-mythos-schwachstellen-software-bsi-100.html (Abruf am 23.04.2026, 21:53 Uhr).
[438] *ZDF*, „Claude Mythos“ von Anthropic: Neues KI-Modell: BSI zeigt sich besorgt, abrufbar unter https://www.zdfheute.de/politik/deutschland/ki-anthropic-claude-mythos-schwachstellen-software-bsi-100.html (Abruf am 23.04.2026, 21:53 Uhr).
[439] *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 21, 25f.
[440] HK-KI-VO/Hilgendorf/Härtlein/*Hilgendorf/Härtlein*, 1. Aufl. 2025, KI-VO Art. 9 Rn. 2.
[441] Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 9 Rn. 3.

zu berücksichtigen sind; Art. 9 Abs. 5 KI-VO zeigt verschiedene Aspekte auf, die anbieterseitig bei der Risikobewältigung zu berücksichtigen sind, damit das verbleibende Risiko des Hochrisiko-KI-Systems auf ein vertretbares Minimum reduziert wird.[442] Mit den nach Art. 9 Abs. 6 KI-VO verpflichtenden Tests soll der Einklang des Systems mit seiner Zweckbestimmung und den Hochrisiko-KI-System-Anforderungen sichergestellt werden und die bestgeeignetsten, gezielten Risikomaßnahmen in Erfahrung gebracht werden, wobei die Tests vor Inverkehrbringen oder Inbetriebnahme zwingend (sowie zu jedem geeigneten Zeitpunkt während des gesamten Entwicklungsprozesses) und nach der Maßgabe geeigneter Testparameter zu erfolgen haben, Art. 9 Abs. 8 KI-VO, sowie unter Realbedingungen stattfinden können, Art. 9 Abs. 7 KI-VO.[443] Im Rahmen des Risikomanagementsystems sind auch nachteilige Auswirkungen auf Personen unter 18 Jahren oder ggf. andere schutzbedürftige Gruppen im Blick zu behalten gem. Art. 9 Abs. 9 KI-VO. Der gesamte Vorgang ist zu dokumentieren.[444]

Im Rahmen der Risikoermittlung und -analyse sind sowohl Risiken der zweckbestimmten Verwendung als auch der vernünftigerweise vorhersehbaren Fehlanwendung (z. B. die ungeprüfte Übernahme eines Entscheidungsvorschlags) zu berücksichtigen, wobei die Vorhersehbarkeit anhand eines verständigen Durchschnittsbeobachters zu ermitteln ist, jedoch subjektive Stärken den Sorgfaltsmaßstab individuell verschärfen können (vgl. auch § 276 BGB, einfache Fahrlässigkeit).[445] Die Risikoermittlung ist auf bestehende Techniken sowie Methoden zu gründen, etwa Besichtigungen oder Begehungen, Kreativitätspraktiken, Checklisten, Analysen von Dokumenten, Unternehmen und Umwelt, Früherkennungssysteme sowie System- und Prozessanalysen, oder Risikoaxonomien, das Heranziehen einschlägiger Szenarioanalysen und Incident Databases.[446]

Es sollten allgemeine **Risikomanagementprinzipien und -prozesse** KI-risikenspezifisch adaptiert werden.[447] Die Art und Weise der Risikoabschätzung ist nicht seitens der KI-VO vorgegeben, es ist auf bereits bekannte Techniken, wie Bayes’sche Netze oder Einflussdiagramme zurückzugreifen, um die Wahrscheinlichkeit des Schadenseintritts und die Schadensintensität zu bewerten (siehe zu Methoden der Risikoanalyse zudem bereits Abschnitt C.I.3).[448] Bei der Risikoabschätzung und -bewertung sind auch vernünftigerweise vorhersehbare Fehlanwendungen, einschließlich menschlichen Verhaltens (unbeabsichtigt sowie böswillig [z. B. Nutzung für einen verbotenen Zweck nach Art. 5 KI-VO, durch Ausnutzung der Funktionspotentiale eines KI-Systems in gezielt täuschender Weise zur Verhaltensveränderung einer Person]) oder Interaktionen mit anderen Systemen zu beachten.[449] Die zu ergreifenden Risikomanagementmaßnahmen zur Reduzierung der ermittelten und bewerteten Risiken beinhalten nicht nur organisatorische Maßnahmen, sondern insbesondere konkrete Maßnahmen bei der Entwicklung und Konzeption des Systems; gem. Art. 9 Abs. 3 KI-VO sind jedoch nur die Risiken, „*die*

[442] Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 9 Rn. 3.
[443] Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 9 Rn. 4.
[444] Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 9 Rn. 17.
[445] BeckOK KI-Recht/*Gerdemann*, 7. Ed. 1.8.2026, KI-VO Art. 9 Rn. 39; Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 9 Rn. 22f.
[446] Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 9 Rn. 24 m. H. a. weitere Instrumente; BeckOK KI-Recht/*Gerdemann*, 7. Ed. 1.8.2026, KI-VO Art. 9 Rn. 35 m. w. N.
[447] BeckOK KI-Recht/*Gerdemann*, 7. Ed. 1.8.2026, KI-VO Art. 9 Rn. 35 m. H. a. weitere Maßnahmen.
[448] Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 9 Rn. 25.
[449] BeckOK KI-Recht/*Gerdemann*, 7. Ed. 1.8.2026, KI-VO Art. 9 Rn. 46; Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 9 Rn. 26.

durch die Entwicklung oder Konzeption des Hochrisiko-KI-Systems oder durch die Bereitstellung ausreichender technischer Informationen angemessen gemindert oder behoben werden können“ zu betrachten, also solche, denen mit den Maßnahmen nach Art. 9 Abs. 5 UAbs. 2 lit. a–c KI-VO begegnet werden kann, wobei nur eine angemessene Risikominderung/-behebung erfolgen muss.[450] Die Auswirkungen und Wechselwirkungen beim Ergreifen der Maßnahmen mit anderen Maßnahmen der Art. 8 ff. KI-VO sind gem. Art. 9 Abs. 4 KI-VO zu beachten; es bleibt den Anbietern offen, flankierende Risikomanagementmaßnahmen zu treffen, zur Reduzierung atypischer Risiken, die nicht bereits durch die Maßnahmen nach Art. 10 ff. KI-VO behandelt sind.[451] Art. 9 Abs. 5 UAbs. 2 lit. a-c KI-VO beinhaltet ein kaskadenartiges Konzept, in dem unterschiedliche Risikominimierungsstufen dem Grundsatz „Konstruktion vor Instruktion“ entsprechend aufeinander folgen, beginnend mit der Wahl einer sicheren Konstruktion, bei deren Unmöglichkeit können Schutzmaßnahmen vorgesehen werden und für verbleibende Restrisiken (die ausdrücklich zugelassen sind) kann der Verweis auf eine Instruktion der Betreiber erfolgen.[452]

In der Praxis sollten – soweit möglich – die Risiken unmittelbar technisch aufgelöst werden; eine frühe sorgfältige Risikoanalyse ist hierfür Voraussetzung, ansonsten sollte für später erkannte Risiken „Re-Engineering“ – eine ingenieursmäßige Neugestaltung von Strukturen und Systemen bzw. deren Ersetzung – verfolgt werden.[453] So wäre bei einer automatisierten KI-Fräse von dieser zu erkennen, dass sich eine menschliche Hand im Fräsen-Bereich befindet und in einem solchen Falle ihr Betrieb zu unterbrechen; ist dies nicht möglich oder die Funktion nicht hinreichend, wären sodann etwa Lichtschranken zur Zielerreichung zu implementieren.[454] Sollte dies nicht hinreichend erfolgreich sein, sollten Minderungs- und Kontrollmaßnahmen (z. B. eine Kameraüberwachung oder ein Not-Aus-Knopf) implementiert und stets eine geeignete Anleitung erstellt werden, damit die Risiken im Einsatz weitestmöglich adressiert werden können; die aus dem Produkthaftungsrecht bekannte Instruktionspflicht kann hierfür als Orientierungspunkt genommen werden.[455]

Am Beispiel eines Übersetzungssystems, das toxische Sprache ausgibt, wären primär Systemanpassungen an diesem vorzunehmen, sodass das System dies nicht mehr tut, wenn dies etwa nicht ausreicht, Maßnahmen zur Inhaltserkennung (z. B. Sicherheitsfilter, die toxische Sprache aussondern) zu implementieren und bei verbleibenden Restrisiken über z. B. die Gefahren toxischer Sprache sowie die weiteren (zuvor genannten) Maßnahmen zu informieren.[456] Diese Information über die Restrisiken und die getroffenen Minderungs- sowie Kontrollmaßnahmen ist in einer (digitalen) Gebrauchsanweisung präzise, korrekt und vollständig darzustellen (siehe Art. 13 KI-VO).[457] Gem. Art. 9 Abs. 5 UAbs. 1 KI-VO sind

---

[450] Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 9 Rn. 30.
[451] Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 9 Rn. 31f.
[452] Hilgendorf/Roth-Isigkeit KI-VO/*Spindler/Gerdemann*, 2. Aufl. 2025, § 5 Rn. 13; Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 9 Rn. 35; HK-KI-VO/Hilgendorf/Härtlein/*Hilgendorf/Härtlein*, 1. Aufl. 2025, KI-VO Art. 9 Rn. 13.
[453] Förster, Die KI-Verordnung in der Praxis, 1. Aufl. 2025, § 2 Rn. 52; zur Begrifflichkeit des Re-Engineering siehe *Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, DFKI*, Re-Engineering, abrufbar unter https://werner.dfki.de/technologydetail/2e453901-e173-477e-bbe3-63328323e349 (Abruf am 09.09.2026, 20:57 Uhr).
[454] Dieses Beispiel ist angelehnt an das Beispiel eines KI-gesteuerten Lasercutters zum Schneiden von Aluminiumplatten nach Förster, Die KI-Verordnung in der Praxis, 1. Aufl. 2025, § 2 Rn. 52.
[455] Förster, Die KI-Verordnung in der Praxis, 1. Aufl. 2025, § 2 Rn. 52.
[456] Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 9 Rn. 36ff.
[457] Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 9 Rn. 39.

Risikomanagementmaßnahmen so zu gestalten, dass sowohl jedes mit einer bestimmten Gefahr verbundene relevante Restrisiko als auch das Gesamtrestrisiko als vertretbar beurteilt werden kann. Eine Schulungspflicht als Risikomanagementmaßnahme zur Bewältigung bzw. Reduzierung von Risiken, die vom Hochrisiko-KI-System für Sicherheit, Grundrechte oder Gesundheit ausgehen können, ist in Art. 9 Abs. 5 KI-VO nicht vorgesehen.[458] Dies spricht indes nicht dagegen, eine solche vorzusehen. Die Vertretbarkeit – Art. 9 KI-VO erwartet kein fehlerfreies oder risikoloses KI-System, sondern nur KI-Systeme mit einem vertretbaren Restrisiko – ist nicht in der KI-Verordnung selbst spezifiziert; solange eine solche Vertretbarkeitsschwelle nicht harmonisiert normiert ist, kann auf bekannte Risikomanagement-Standards zurückgegriffen werden, wie etwa den ISO/IEC Guide 51:2014, Ziff. 6.2.1, mit den darin vorgeschlagenen zu berücksichtigenden Faktoren, wie die aktuellen Wertvorstellungen der Gesellschaft oder der Suche nach einem bestmöglichen Ausgleich zwischen dem Ideal absoluter Sicherheit und dem tatsächlich Erreichbaren.[459] Die gem. Art. 9 Abs. 6 S. 1 KI-VO verpflichtenden Tests – unter denen eine Reihe von Aktivitäten zur Identifizierung sowie Evaluierung mit Eigenschaften des Testgegenstands verstanden werden können[460] – gem. Art. 9 Abs. 6–8 KI-VO des KI-Systems sind ein wesentlicher Bestandteil des Risikomanagementsystems und von besonderer praktischer Bedeutung, da sich KI-Systeme durch besondere Komplexität und Variabilität ausweisen, die Prognosen erschweren; KI-Systeme verhalten sich oft anders bei Unterschieden zwischen **Trainingslandschaft und praktischer Anwendungsumgebung**.[461] Dies ist insbesondere bei Systemen maschinellen Lernens von Bedeutung, bei denen bei Design und Durchführung der Tests die bekannten Probleme der Systeme maßgeblich zu berücksichtigen sind, wie z. B. sogenannte „Dataset Shifts“, die vorliegen, wenn die gemeinsame Verteilung von Ein- und Ausgabedaten zwischen der Trainings- sowie der Testphase differiert.[462] Für die Tests können und sollten professionalisierte Anbieter hinzugezogen werden.[463] Hinreichende Testungen sind insbesondere mit Blick auf Haftungsfragen von großer Bedeutung.[464] Mit Blick auf einen beschleunigten Entwicklungsprozess und Wirklichkeitsnähe sind insbesondere Tests in **Reallaboren**, Art. 9 Abs. 7 KI-VO, vorteilhaft.[465] Das System ist gem. Art. 9 Abs. 8 S. 1 KI-VO jedenfalls vor dem Inverkehrbringen oder der Inbetriebnahme, also mindestens einmal

[458] Hilgendorf/Roth-Isigkeit KI-VO/*Spindler/Gerdemann*, 2. Aufl. 2025, § 5 Rn. 12, 14.
[459] Hilgendorf/Roth-Isigkeit KI-VO/*Spindler/Gerdemann*, 2. Aufl. 2025, § 5 Rn. 15ff.
[460] BeckOK KI-Recht/*Gerdemann*, 7. Ed. 1.8.2026, KI-VO Art. 9 Rn. 87 m. w. H.
[461] Förster, Die KI-Verordnung in der Praxis, 1. Aufl. 2025, § 2 Rn. 53; Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 9 Rn. 43; BeckOK KI-Recht/*Gerdemann*, 7. Ed. 1.8.2026, KI-VO Art. 9 Rn. 87f.; zu möglichen Aspekten und Problemen hinsichtlich der Tests im Zusammenhang mit Justiz-KIs siehe auch Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 55ff.
[462] BeckOK KI-Recht/*Gerdemann*, 7. Ed. 1.8.2026, KI-VO Art. 9 Rn. 88f. m. w. N.; hinsichtlich Dataset Shifts siehe etwa *Quiñonero-Candela/Sugiyama/Schwaighofer/Lawrence*, Dataset Shift in Machine Learning, Neural Information Processing Series, 1. Aufl. 2008, abrufbar unter https://direct.mit.edu/books/edited-volume/3841/Dataset-Shift-in-Machine-Learning (Abruf am 10.09.2026, 16:22 Uhr).
[463] BeckOK KI-Recht/*Gerdemann*, 7. Ed. 1.8.2026, KI-VO Art. 9 Rn. 90.
[464] Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 9 Rn. 43.
[465] Vgl. Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 9 Rn. 44; Förster, Die KI-Verordnung in der Praxis, 1. Aufl. 2025, § 2 Rn. 55.

zu testen; wann ein geeigneter Zeitpunkt vorliegt, ist nicht pauschal zu beantworten, wann und wie oft ein Test erfolgt, hängt damit im Wesentlichen vom eigenen Ermessen ab.[466] Mit Blick auf die wesentliche Bedeutung des judikativen Einsatzes sollte aber eine regelmäßige Frequenz gewählt werden, zumindest beispielsweise angeglichen an Updates oder Systemversionen-Änderungen. Welche Metriken und Wahrscheinlichkeitsschwellenwerte im Rahmen von Art. 9 Abs. 8 S. 2 KI-VO anzuwenden sind, ist nicht näher definiert; ihre Angemessenheit ist kontextspezifisch, zudem besteht noch keine empfohlene Vorgehensweise.[467] Zwar sind die Werte vorab festzulegen, eine Anpassung nach dem Test und eine Wiederholung anhand der neuen Zielwerte erscheint jedoch auch im Sinne eines kontinuierlichen, iterativen Risikomanagements möglich.[468]

cc. Genauigkeit, Robustheit und Cybersicherheit

Des Weiteren sollten die Aspekte insbesondere von Art. 15 KI-VO beachtet werden. Dieser steht im engen Zusammenhang mit Art. 9 KI-VO, da er grundlegende Aspekte für das Risikomanagementsystem liefert.[469] Auch steht Art. 15 KI-VO im engen Zusammenhang zu Art. 25 und 32 DSGVO (hierzu siehe noch zum Datenschutz im Abschnitt C.II.5.b.bb), die als sein Gegenstück und – als im Gegensatz zu ihm auch organisatorische Maßnahmen erfassend – mit Vorbildfunktion ausgestattet angesehen werden.[470] Nach Art. 15 Abs. 1 KI-VO müssen Hochrisiko-KI-Systeme *„so konzipiert und entwickelt [werden], dass sie ein angemessenes Maß an Genauigkeit, Robustheit und Cybersicherheit erreichen und in dieser Hinsicht während ihres gesamten Lebenszyklus beständig funktionieren“.* Als an der Konzeption und Entwicklung anknüpfend, sind die Aspekte bereits frühestmöglich zu beachten und fortlaufend sicherzustellen.[471]
Während die **Genauigkeits- und Robustheitsvoraussetzungen** die **Außenwelt vor Systemfehlern schützen** sollen, soll die **Cybersicherheit das System vor Zugriffen von außen** schützen.[472] Das angemessene Maß ist nicht abstrakt, sondern anhand des konkreten Anwendungszwecks zu bestimmen, auch sind die Aspekte bei gegenseitiger Beeinflussung in ein angemessenes Verhältnis untereinander zu bringen.[473]
Ausgerichtet anhand des Verordnungszwecks sollte ein Niveau angestrebt werden, sodass die implementierten Vorkehrungen einen hinreichenden Schutz unter Beachtung eines Kosten-Nutzen- bzw. Kosten-Risiko-Verhältnisses für die gem. Art. 1 Abs. 1 KI-VO aufgeführten Aspekte sicherstellen.[474] Dementsprechend ist auch die Genauigkeit (die nicht definiert wird; sie kann jedoch in etwa als die Fähigkeit des Systems begriffen werden, Sachverhalte korrekt zu beurteilen) relativ zu verstehen, die jedoch nach Erwägungsgrund 74 S. 1 der KI-VO anhand von der Zweckbestimmung des Systems als auch anhand des Standes der Technik zu bestimmen ist; je höher das Risiko für ein geschütztes Gut bzw. je weiter der Stand

[466] BeckOK KI-Recht/*Gerdemann*, 7. Ed. 1.8.2026, KI-VO Art. 9 Rn. 95ff.; Förster, Die KI-Verordnung in der Praxis, 1. Aufl. 2025, § 2 Rn. 55.
[467] Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 9 Rn. 46.
[468] Bomhard/Pieper/Wende/*König*, 1. Aufl. 2025, KI-VO Art. 9 Rn. 48.
[469] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 15 Rn. 9.
[470] BeckOK KI-Recht/*Buchner*, 7. Ed. 1.8.2026, KI-VO Art. 15 Rn. 29.
[471] BeckOK KI-Recht/*Buchner*, 7. Ed. 1.8.2026, KI-VO Art. 15 Rn. 30.
[472] HK-KI-VO/Hilgendorf/Härtlein/*Hilgendorf/Härtlein*, 1. Aufl. 2025, KI-VO Art. 15 Rn. 1.
[473] Bomhard/Pieper/Wende/*Moritz*, 1. Aufl. 2025, KI-VO Art. 15 Rn. 33.
[474] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 15 Rn. 30.

der Technik ist, desto höher sind die Anforderungen an die Genauigkeit zu stellen.[475] Die angemessene Genauigkeit ist elementar für eine vertrauenswürdige KI.[476] Mit Blick auf den judikativen Einsatz ist daher höchste Genauigkeit eine elementare Voraussetzung. Die Genauigkeitsmaße und -metriken sind anbieterseitig in der Betriebsanleitung anzugeben, Art. 15 Abs. 3 und Art. 13 Abs. 2 KI-VO; sie sollten in klarer, leicht verständlicher Weise ohne Missverständnisse und ohne irreführende Aussagen zur Verfügung gestellt werden, Erwägungsgrund 74 der KI-VO.[477]
Die Robustheit ist ebenfalls nicht definiert, lässt sich aber anhand von Art. 15 Abs. 4 UAbs. 1 KI-VO und Erwägungsgrund 75 der KI-VO konturieren und in der Folge verstehen als Widerstandsfähigkeit gegenüber Fehlern, Störungen oder Unstimmigkeiten, die innerhalb eines KI-Systems oder der Umgebung, in der das System betrieben wird, auftreten können; erfasst sind insbesondere auch solche Fehlerquellen, die aus Interaktionen natürlicher Personen mit dem KI-System oder anderer Systeme mit diesem resultieren (bspw. Netzstörungen oder systemische Effekte wegen Störungen von Intermediären im Rahmen des Datenaustauschs, Inkonsistenzen in den Eingabedaten oder widersprüchliche Informationen, „Feedback“-Schleifen oder menschliche Fehler).[478] Bis zum Erlass harmonisierter KI-Normen bzw. bis zum Vorliegen der Benchmarks sowie Messmethoden nach Art. 15 Abs. 2 KI-VO sollte auf relevante Normungen, wie etwa ISO/IEC TR 24029-1:2021 Rückgriff genommen werden.[479] Für die Robustheit sind gem. Art. 15 Abs. 4 UAbs. 1 S. 2 KI-VO technische und organisatorische Maßnahmen zu ergreifen, wobei das System gem. Art. 15 Abs. 4 UAbs. 1 S. 1 KI-VO so widerstandsfähig wie möglich sein soll. Technische Redundanzen werden nach Art. 15 Abs. 4 UAbs. 2 KI-VO als „Königsweg“ zur Erreichung der Robustheit angesehen, da sich mittels dieser mögliche Ausfälle durch andere Systemressourcen kompensieren lassen.[480] Umfasst hiervon sind u. a. regelmäßige Backups, Unterbrechungsmechanismen, bei einer Funktion außerhalb vordefinierter Grenzen (vgl. Erwägungsgrund 75 S. 4 KI-VO) sowie die ausdrücklich in der Norm genannten Sicherungs- und Störungssicherheitspläne, mittels derer erreicht werden soll, dass trotz Störungen keine Gefahren ausgelöst werden.[481] Hierbei werden vorab unterschiedliche Fehlerszenarien umfassend betrachtet und für den jeweiligen Fall spezifische Reaktionsstrategien definiert; auch wird festgelegt, welche Störungen systemseitig erkannt werden, welche Eskalationskaskaden vorgesehen werden sowie die Reihenfolge der konkreten Schutzmaßnahmen.[482] Die Pläne können etwa programmierte Abschaltfunktionalitäten beinhalten, die ein Hochrisiko-KI-System abschalten oder in einen Ursprungszustand zurückversetzen.[483] Im Kontext der technischen Robustheit werden seitens der KI-Expertengruppe auch sog. Auffangpläne vorgeschlagen, sollte ein System unerwarteten Situationen oder feindlichen Angriffen ausgesetzt sein, die als Sicherheitsvorkehrungen von Systemen maschinellen Lernens auf regelbasierte umstellen oder einen menschlichen Eingriff anfordern, wobei freilich auch Szenarien bedacht werden sollten, in denen kein

[475] BeckOK KI-Recht/*Buchner*, 7. Ed. 1.8.2026, KI-VO Art. 15 Rn. 32f.
[476] BeckOK KI-Recht/*Buchner*, 7. Ed. 1.8.2026, KI-VO Art. 15 Rn. 31.
[477] Förster, Die KI-Verordnung in der Praxis, 1. Aufl. 2025, § 2 Rn. 105.
[478] BeckOK KI-Recht/*Buchner*, 7. Ed. 1.8.2026, KI-VO Art. 15 Rn. 34; Hilgendorf/Roth-Isigkeit KI-VO/*Spindler/Gerdemann*, 2. Aufl. 2025, § 5 Rn. 58; Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 15 Rn. 52.
[479] Vgl. Hilgendorf/Roth-Isigkeit KI-VO/*Spindler/Gerdemann*, 2. Aufl. 2025, § 5 Rn. 57ff.
[480] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 15 Rn. 54.
[481] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 15 Rn. 55.
[482] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 15 Rn. 55.
[483] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 15 Rn. 55.

unmittelbarer menschlicher Eingriff möglich ist.[484] Trivialere Maßnahmenkomplexe, wie etwa Benutzerschnittstellen-Eingrenzungen, um Manipulationen des Codes zu verhindern, die Trennung von Datensphären sowie die Ausgabe von verständlichen Fehlermeldungen sollten freilich auch ergriffen werden.[485] Gemäß Art. 15 Abs. 4 UAbs. 3 KI-VO sind bei selbstlernenden Systemen (maschinellen Lernens) sog. Rückkopplungsschleifen (normseitig definiert als „*das Risiko möglicherweise verzerrter Ausgaben, die künftige Vorgänge beeinflussen*“) zu beseitigen oder so gering wie möglich zu halten. Auch soll sichergestellt werden, dass auf sie mit geeigneten Risikominderungsmaßnahmen eingegangen wird. Solche Rückkopplungsschleifen können bestehende Diskriminierungen fortschreiben und verstärken, vgl. Erwägungsgrund 67 S. 7 der KI-VO.
Technische risikomindernde Maßnahmen gegen Rückkopplungsschleifen beinhalten u. a. die sorgfältige, dem Training vorgelagerte, Datenauswahl und -aufbereitung, kontinuierliches Monitoring der Systemausgaben hinsichtlich Verzerrungen sowie Retraining mit unbefangenen Daten in regelmäßigen Abständen.[486]
Art. 15 Abs. 5 UAbs. 1 KI-VO versteht Cybersicherheit derart, dass Hochrisiko-KI-Systeme **widerstandsfähig gegen Versuche unbefugter Dritter** sein müssen, ihre Verwendung, Ausgaben oder Leistung durch Ausnutzung von Systemschwachstellen zu verändern.[487] Erfasst sind über Schwachstellen, die KI-spezifisch sind, hinaus, jegliche solche in Hardware, Betriebssystem und Netzwerkkomponenten.[488] Der Anreiz, Systeme anzugreifen, steigt, je sensibler deren Ausgaben sind.[489] Mit Blick auf den judikativen Einsatz ist daher mit einer erheblichen Bedrohungslage zu rechnen. Die Angriffsmethoden hängen wesentlich von den Fähigkeiten des Angreifers und der bestehenden Kontrollreichweite bezüglich der angegriffenen KI ab; es bestehen verschiedene KI-spezifische Angriffstaktiken, etwa anknüpfend an die Trainingsdaten, das Modell in seiner Trainings- oder Einsatzumgebung; passend dazu nennt Art. 15 Abs. 5 UAbs. 3 KI-VO das „data-poisoning“, das „model-poisoning“, „adversarial examples“, „model evasions“, Angriffe auf vertrauliche Daten und Modellmängel.[490] Nach Art. 15 Abs. 5 UAbs. 2 KI-VO müssen die technischen Lösungen zur Gewährleistung der Cybersicherheit von Hochrisiko-KI-Systemen den jeweiligen Umständen und Risiken angemessen sein. Damit wird eine Angemessenheit, im Sinne eines Standhaltens einer Kosten-Nutzen-Abwägung verlangt; die spezifischen Risiken sind (auf Grundlage einer Risikosicherheitsbewertung nach Art. 9 KI-VO) zu ermitteln.[491]
Es sind nur technische Lösungen vorgeschrieben, gleichwohl sind auch organisatorische Maßnahmen, wie etwa Kontrollen der Datenqualität wirkungsvoll und empfehlenswert.[492] Art. 15 Abs. 5 UAbs. 3 KI-VO zählt exemplarisch Cybersicherheitsrisiken auf, die spezifisch für Systeme maschinellen Lernens auf Attacken sind und die ggf. durch Maßnahmen verhütet, erkannt, auf die reagiert, beseitigt und kontrolliert werden sollen.[493] Hierbei sollten u. a. Sicherheitskontrollen, Vorfallbewältigungspläne, die festlegen, wie im Angriffsfalle verfahren werden soll,

[484] BeckOK KI-Recht/*Buchner*, 7. Ed. 1.8.2026, KI-VO Art. 15 Rn. 50.
[485] Marly SoftwareR-HdB/*Beurskens*, 8. Aufl. 2024, § 16 Rn. 62.
[486] Bomhard/Pieper/Wende/*Moritz*, 1. Aufl. 2025, KI-VO Art. 15 Rn. 58.
[487] BeckOK KI-Recht/*Buchner*, 7. Ed. 1.8.2026, KI-VO Art. 15 Rn. 55, 36.
[488] BeckOK KI-Recht/*Buchner*, 7. Ed. 1.8.2026, KI-VO Art. 15 Rn. 56.
[489] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 15 Rn. 59.
[490] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 15 Rn. 63f.
[491] BeckOK KI-Recht/*Buchner*, 7. Ed. 1.8.2026, KI-VO Art. 15 Rn. 57f.; Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 15 Rn. 67.
[492] Vgl. Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 15 Rn. 66.
[493] Vgl. Hilgendorf/Roth-Isigkeit KI-VO/*Spindler/Gerdemann*, 2. Aufl. 2025, § 5 Rn. 61.

implementiert werden und besonderes Augenmerk auf Schwachstellen gelegt werden, die mehreren KI-Systemen immanent sind.[494] Maßnahmen gegen „data-poisoning“ verhindern diese bestenfalls oder offenbaren diese zur Behebung und Kontrolle und stellen idealiter sicher, dass nur geeignete Daten zum Training herangezogen werden.[495] Datenmanagement (siehe auch Art. 10 KI-VO) sowie Regeln für die Datensammlung und -verwertung, sog. Anomalie-Sensoren, die Abweichungen bei der Datenverteilung erkennen, Daten- und Ergebnis-Validierung sowie die Verwendung robuster Modelle können als Maßnahmen implementiert werden.[496] „Adversarial examples“ oder „model evasions“ sind Eingabedaten, die das KI-Modell zu Fehlern verleiten sollen und als Angriffsmethoden für den Menschen oftmals schwerlich erkennbar und in der Folge besonders gefahrträchtig, insbesondere, je minimaler und unscheinbarer ihre Störung ist; sie haben u. a. gezielt (falsche) Ausgaben oder Fehlfunktionen des Systems zur Folge.[497] Risiken können in diesem Zusammenhang über die Offenlegung von Modellinformationen aufkommen; das Zusammenfassen mehrerer quelloffener KI-Modelle zu einem sog. „Ensemble“, die Kennzeichnung vertrauenswürdiger Eingabedaten mit einem digitalen Wasserzeichen, der Einsatz robusterer Modelle und weitere Methoden, wie etwa defensive Destillation, können Risiken mildern.[498] Mittels der Entfernung personenbezogener Daten und Datensatzbereinigung kann Vertraulichkeitsangriffen entgegengewirkt werden.[499] In der Praxis sollte beachtet werden, dass nach Art. 42 Abs. 2 KI-VO Hochrisiko-KI-Systeme, die im Rahmen eines der Cybersicherheitszertifizierungssysteme, die gemäß der Verordnung (EU) 2019/881 (sog. EU Cybersecurity Act), deren Fundstellen im Amtsblatt der Europäischen Union veröffentlicht wurden, zertifiziert wurden oder für die eine solche Konformitätserklärung erstellt wurde, die Vermutung gilt, dass sie die in Art. 15 KI-VO festgelegten Cybersicherheitsanforderungen erfüllen, sofern diese Anforderungen von der Cybersicherheitszertifizierung oder der Konformitätserklärung oder Teilen davon abgedeckt sind. Derartige Harmonisierungsmöglichkeiten zur Anforderungserfüllung sollten in Betracht gezogen werden.

dd. Systemseitige oder sonstige Vorkehrungen gegen unzulässigen Einfluss Dritter auf Entwicklung, Gestaltung und Kontrolle des Systems

Informationstechnische Systeme werden zumeist von Drittanbietern bezogen, womit einhergeht, dass diese wesentliche Einflussmöglichkeiten auf die Entwicklung, Gestaltung und Kontrolle des Systems besitzen; die Vermeidung unzulässiger Einflussnahme ist somit schon bei der **Auswahl von Kooperationspartnern** von Bedeutung.[500] Drittanbieter stützen sich wiederum oftmals auf Bestandteile von weiteren Dritten, sodass eine **KI-Lieferkette** entsteht, bei der der Endkonsument – hier die Judikative – zumeist nicht über einen hinreichenden Einblick in bzw. eine hinreichende Expertise für eine umfassende Risikobetrachtung und -adressierung in Bezug auf das bezogene KI-System verfügt.[501]

[494] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 15 Rn. 75.
[495] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 15 Rn. 79.
[496] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 15 Rn. 79.
[497] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 15 Rn. 83f.
[498] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 15 Rn. 87f.
[499] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 15 Rn. 91.
[500] Vgl. *Mulligan/Bamberger*, Procurement as Policy: Administrative Process for Machine Learning, Berkeley Technology Law Journal 34 (2019), 773, 779ff.
[501] *National Cyber Security Centre et al.*, Guidelines for secure AI system development, abrufbar unter https://www.ncsc.gov.uk/sites/default/files/documents/Guidelines-for-secure-AI-system-development.pdf (Abruf am 23.04.2026, 13:14 Uhr), S. 7; zu beachten ist zudem, dass die meisten KI-

Eine erste Handlungsoption besteht daher darin, die unmittelbaren Bezugspartner des informationstechnischen Systems dazu zu verpflichten, die seitens ihnen von weiteren Dritten bezogenen Bestandteile auf die für die Justiz notwendigen Sicherheitsstandards zu überprüfen und hierfür mit ihnen zu kooperieren sowie eigene Sicherheitsvorkehrungen zu implementieren.[502] Dies sollte sich in regelmäßigen Berichten vorgelegt werden lassen und die Manipulationsfreiheit – sowohl vor Implementierung des informationstechnischen Systems (oder von Updates) als auch danach – von externen, vertrauenswürdigen Institutionen bzw. justizeigenen Expertengremien überprüft werden.[503] Auch mit Blick auf diese Überprüfung können KI-Systeme unterstützend eingesetzt werden.[504] Denkbar wäre darüber hinaus der Ansatz vollständiger technischer Transparenz etwa mittels open source Codes verfolgbar.[505] Die Informationssicherheit ist gerade im judikativen Bereich von elementarer Bedeutung; Drittanbieter sowie die Lieferketten-Thematik sollten daher Berücksichtigung als eigener Risikofaktor im Rahmen der Risikobewertung und im Sicherheitskonzept finden.[506] Es sollte neben der Inanspruchnahme von Zertifizierungsmechanismen eine fortlaufende Auditierung für das informationstechnische System über den ganzen Lebenszyklus hinweg implementiert werden.[507] Hinsichtlich der Reichweite der Zertifizierung bestehen unterschiedliche Ansichten. Im Kontext regelbasierter Systeme schlägt hierbei *Rollberg* eine Zertifizierung allein im Hinblick darauf vor, dass die in einer Anleitung zum System enthaltenen Angaben der Funktionsweise entsprechen.[508] Die Untersuchung auf Diskriminierungen sowie eine „Rechtmäßigkeitsprüfung“ mit Blick darauf, ob bzw. welche Rechtsansichten im System (z. B. neben der höchstrichterlichen Ansicht auch abweichende vorinstanzliche Ansichten oder Literaturansichten) enthalten sind, werden als zu weitgehende inhaltliche Einflussnahme im Rahmen der richterlichen Unabhängigkeit abgelehnt.[509] *Biallaß* stellt hingegen die Vorteilhaftigkeit für die richterliche Unabhängigkeit heraus, durch eine Prüfung, ob sämtliche

---

Tools keine eigenen KI-Modelle haben, sondern als sog. KI-Wrapper um KI-Modelle anderer Anbieter herumgebaut sind, *Kuhlmann*, NJW 2026 2774, 2775.

[502] *National Cyber Security Centre et al.*, Guidelines for secure AI system development, abrufbar unter https://www.ncsc.gov.uk/sites/default/files/documents/Guidelines-for-secure-AI-system-development.pdf (Abruf am 23.04.2026, 13:14 Uhr), S. 7; siehe auch *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 21.

[503] *European Commission for the Efficiency of Justice (CEPEJ),* European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment, abrufbar unter https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c (Abruf am 24.04.2026, 14:58 Uhr), S. 9ff.; *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 340, 373.

[504] Vgl. mit Blick auf Diskriminierungskontexte *European Commission for the Efficiency of Justice (CEPEJ),* European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment, abrufbar unter https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c (Abruf am 24.04.2026, 14:58 Uhr), S. 9.

[505] *European Commission for the Efficiency of Justice (CEPEJ),* European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment, abrufbar unter https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c (Abruf am 24.04.2026, 14:58 Uhr), S. 11.

[506] Vgl. *National Cyber Security Centre et al.*, Guidelines for secure AI system development, abrufbar unter https://www.ncsc.gov.uk/sites/default/files/documents/Guidelines-for-secure-AI-system-development.pdf (Abruf am 23.04.2026, 13:14 Uhr), S. 7, 9ff.

[507] *Martini*, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz, 1. Aufl. 2019, S. 323f.; *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 340, 373f.

[508] *Rollberg*, Algorithmen in der Justiz, 1. Aufl. 2020, S. 111ff.

[509] *Rollberg*, Algorithmen in der Justiz, 1. Aufl. 2020, S. 108ff.; *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 292f.

Auslegungsvarianten systemseitig abgebildet werden; die rechtliche Würdigung bei Abbildung aller Varianten bliebe allein beim menschlichen Richter.[510] Auch hinsichtlich einer Überprüfung seitens einer Zertifizierung im Hinblick darauf, dass es durch den Einsatz des Systems nicht zu einer Diskriminierung von bestimmten Personengruppen komme, sei nichts einzuwenden.[511] Mit Blick auf die Bias-Problematik ist dem Ansatz von *Biallaß* beizupflichten. Es steht mit Blick auf die Anzeige von Auslegungsvarianten zu befürchten, dass selbst, wenn keine Überprüfung durch die Zertifizierungsstelle stattfindet und auch wenn es die Aufgabe des Gerichts bleibt, sämtliche Auslegungsvarianten im Blick zu behalten[512], dass die systemseitig angezeigten Auslegungsvarianten präsumtiv öfter gewählt werden als nicht angezeigte. So könnte beispielsweise das bereitstellende Unternehmen versucht sein, daher bei Rechtsfragen, die Unternehmen betreffen, nur (die) für diese förderliche(n) Ansicht(en) darzustellen. Zur Veranschaulichung könnte auch beispielhaft der Fall divergierender obergerichtlicher Rechtsprechung angeführt werden. Die z. B. im System singulär angezeigte eine obergerichtliche Ansicht könnte somit durch ihre alleinige Aufführung öfter bei der Entscheidungsfindung in anderen Fällen zugrunde gelegt werden, wodurch ihr faktisch eine übermäßige Bedeutung verliehen würde, unabhängig von ihrer (ggf. nicht bestehenden) Vorzugswürdigkeit gegenüber der anderen obergerichtlichen Ansicht. Die Tatsache, dass es dennoch Aufgabe des Gerichts bleibt, sich aller Auslegungsvarianten bewusst zu bleiben und den Maßstab zu ermitteln, ab wann eine Diskriminierung vorliegt, bleibt unabhängig davon bestehen, die Anwendung durch den Richter wird nur vorab auf eine „objektivere“/breitere Grundlage gestellt. Durch die Ausgabe aller Ansichten und der Zertifizierung kann somit verhindert werden, dass einzelnen Ansichten eine faktisch übermäßige Bedeutung verliehen wird. Der Gefahr nicht berücksichtigter Ansichten und einem (dadurch) fälschlichen, übermäßigen Vertrauen in die Zertifizierung kann hingegen durch entsprechende Hinweise im System begegnet werden. Hinsichtlich der Zertifizierungsfrequenz, also wann eine neue Zertifizierung vorzunehmen ist, kann eine Zertifizierung bei jedem neuen Update oder jeder Programmversion gewählt werden.[513] Mit Blick auf die aufzuwendenden Ressourcen für eine Zertifizierung kann auch die Option gewählt werden, bestimmte Grenzen zu definieren, ab denen eine neue Zertifizierung notwendig wird.[514]
Sollten Rechtsstreitigkeiten zu entscheiden sein, bei denen die Rechtsperson, die das informationstechnische System entwickelt oder betreibt, selbst Partei ist oder (offensichtlich) eigene Interessen am Ausgang des Verfahrens hat, ist ausnahmsweise die Unparteilichkeit der Richterinnen und Richter verstärkt in den Blick zu nehmen.[515] In derartigen Fallgestaltungen sollte angedacht werden, das zur Verfügung gestellte informationstechnische System nicht zu benutzen, um dem Anreiz unzulässiger Einflussnahme[516] vorzubeugen. Mit Blick auf den zivilprozessualen Bereich dürfte dies weitestgehend unproblematisch sein. Im Hinblick auf die Gerichtsbarkeiten, in denen der Staat (in Form der Exekutive) regelmäßig Prozessbeteiligter ist (z. B. bei der Verwaltungsgerichtsbarkeit), sollte hingegen beachtet werden, dass bei Ansiedelung der informationstechnischen Systeme in der Hand der Exekutive, diese regelmäßig

[510] *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 293.
[511] *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 293.
[512] *Rollberg*, Algorithmen in der Justiz, 1. Aufl. 2020, S. 111.
[513] *Rollberg*, Algorithmen in der Justiz, 1. Aufl. 2020, S. 113.
[514] Vgl. *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 293.
[515] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 168, 246f.
[516] Vgl. *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 168, 246f.

eigene Interessen am Verfahrensausgang hat (hierbei ist freilich von Relevanz, bei welcher juristischen Person das informationstechnische System angesiedelt ist). Insoweit zeigt sich in derartigen Bereichen in besonderem Maße, dass die Ansiedelung des informationstechnischen Bereichs bei der Judikative der „Königsweg“[517] ist.

c. Zusammenfassung

Die informationstechnischen Systeme müssen sowohl im entscheidungsunterstützenden als auch im entscheidungsersetzenden Einsatz vor internen und externen Störungen (z. B. Softwarefehler, Stromausfall, Hackerangriff) geschützt werden. Dritte dürfen keinen unzulässigen unmittelbaren oder mittelbaren Einfluss auf Entwicklung, Gestaltung und Kontrolle des Systems nehmen. Das eingesetzte System darf sich bei seiner Entscheidungsfindung nur an der Gesetzesbindung orientieren.
Die **Sicherheit des informationstechnischen Systems ist über seinen gesamten Lebenszyklus** zu betrachten, von Design und Entwicklung über die Bereitstellung bis hin zu Betrieb und Wartung. Es sollten eine Risikobewertung für das jeweilige System sowie ein angepasstes Informationssicherheitsmanagementsystem erfolgen. Angesichts der sensiblen Daten und der massiven Auswirkungen einer Systemkorruption sollten höchste Sicherheitsraten angestrebt werden. Der Übergang von der analogen Aktenarbeit zur digitalen Aktenarbeit und nun die mögliche Rechtsprechungsbeeinflussung durch manipulierte informationstechnische Systeme haben die informationstechnische Gefährdungslage erheblich verschärft, weshalb möglichst umfassende Maßnahmen zur Erreichung eines ausreichenden Sicherheitsniveaus zu ergreifen sind. Wesentliches Augenmerk sollten **präventive Absicherungen** mit umfangreichen Testphasen und Evaluierungen vor dem Ausrollen der Systeme erfahren, beispielsweise in Form von öffentlichen Hacking-Tests. Insbesondere sollten auch nur authentifizierte Geräte auf die Systeme zugreifen dürfen. Mit dem Einsatz informationstechnischer Systeme sollte ein **präventives, umfassendes Sicherheits- und Cyber-Resilienzkonzept** einhergehen. KI-spezifisch sind **Prompt-Injections** weitestmöglich abzuwehren, u. a. durch **strukturierte Eingabemasken, Filter und Vorgaben im Systemprompt**. Auch ist das erhebliche Gefährdungspotential durch LLMs, die zunehmend auch als Angriffswerkzeuge benutzt werden, zu beachten.
Unabhängig von der Einordnung als Hochrisiko-KI-System sollte stets ein **Risikomanagementsystem** nach Art. 9 KI-VO als kontinuierlicher und iterativer Prozess eingerichtet werden, um die Gefahr einer Falscheinordnung und die daraus folgenden Konsequenzen auszuschließen. Jedenfalls sollte zumindest eine Bund/Länder-übergreifende Verständigung auf Expertengrundlage zur Auslegung der KI-VO, insbesondere zur Einordnung von Systemen als Hochrisiko-KI, erarbeitet und etabliert werden. Es gilt der Grundsatz „Konstruktion vor Instruktion“, Risiken sollten bestenfalls unmittelbar technisch aufgelöst werden, wobei ein vertretbares Restrisiko des KI-Systems zulässig ist. Das System ist vor der Inbetriebnahme und regelmäßig, zum Beispiel bei Updates oder neuen Systemversionen, zu testen. Hierbei sollten Reallabore und professionelle Anbieter für Tests einbezogen werden.
Zudem sind die Anforderungen an Genauigkeit, Robustheit und Cybersicherheit nach Art. 15 KI-VO zu beachten. Für den Einsatz in der Justiz ist höchste Genauigkeit eine elementare Voraussetzung. Insbesondere technische Redundanzen erscheinen für die Gewährleistung der Robustheit als wesentlich. Gegen KI-spezifische Angriffe wie

[517] *Radke*, jM 2016, 8, 12.

„data-poisoning“, „model-poisoning“ oder „adversarial examples“ sind technische Maßnahmen zu ergreifen, die durch organisatorische Maßnahmen flankiert werden sollten. Harmonisierungsoptionen, wie die Vermutungswirkung gem. Art. 42 Abs. 2 KI-VO einer Zertifizierung nach dem EU Cybersecurity Act, sollten in Betracht gezogen werden.
Wenn informationstechnische Systeme von Drittanbietern bezogen werden, ist die KI-Lieferkette als eigener Risikofaktor zu behandeln. Vorgeschlagen wird insbesondere, die unmittelbaren Bezugspartner zu verpflichten, die von weiteren Dritten bezogenen Bestandteile auf die für die Justiz notwendigen Sicherheitsstandards zu überprüfen und hierüber regelmäßig Bericht zu erstatten. Die Manipulationsfreiheit des informationstechnischen Systems sollte durch externe, vertrauenswürdige Institutionen bzw. justizeigene Expertengremien überprüft werden. Neben Zertifizierungen sollte über den gesamten Lebenszyklus eine **Auditierung** erfolgen. Die Zertifizierung sollte sich darauf erstrecken, dass das System nicht diskriminiert und sämtliche Auslegungsvarianten berücksichtigt und abbildet.
Ist die Rechtsperson, die das informationstechnische System entwickelt oder betreibt, selbst Partei eines zu entscheidenden Rechtsstreits oder hat diese (offensichtlich) eigene Interessen am Ausgang des Verfahrens, sollte angedacht werden, das System nicht zu nutzen. Dies ist insbesondere bei der Ansiedelung des Systems in der Hand der Exekutive relevant im Hinblick auf Gerichtsbarkeiten, bei denen der Staat regelmäßig Prozessbeteiligter ist. In diesen Fällen zeigt sich hervorgehoben die Ansiedelung informationstechnischer Systeme bei der Judikative als „Königsweg“.

4. Transparenz staatlichen Handelns und Information über den Einsatz von Künstlicher Intelligenz

In dem ersten Arbeitspaket (Milestone 1) des Forschungsprojekts „Künstliche Intelligenz und richterliche Entscheidungsfindung“[518] sind weitere, wesentliche rote Linien herausgearbeitet worden, die dem Themenfeld der Transparenz staatlichen Handelns und der Information über den Einsatz von Künstlicher Intelligenz zugeordnet werden können. Diese roten Linien sind im Wesentlichen übergreifender Natur, zeigen jedoch je nach Einsatzszenario unterschiedliche Ausprägungen.

a. Übergreifende rote Linien für jeden Einsatz

Übergreifend, d. h. sowohl für den entscheidungsersetzenden als auch für den entscheidungsunterstützenden Einsatz von informationstechnischen Systemen, gilt, dass der Einsatz den **Transparenzanforderungen** staatlichen Handelns genügen muss. Die Parteien sowie die Öffentlichkeit sind daher unaufgefordert über das **„Ob“** und **„Wie“** des Einsatzes eines informationstechnischen Systems mit Blick auf dessen abstrakte Funktionsweise, d. h. die Arbeitsweise des Systems selbst, und dessen konkretes Einsatzszenario zu **informieren**.[519] Es ist ferner transparent zu machen, auf welche Art und Weise das System zu seiner Entscheidung (**abstrakte Funktionalität**) gekommen ist.[520] Die Parteien müssen die involvierte Logik des Systems nachvollziehen und die wesentlichen Gründe erkennen können, auf denen die Entscheidung des Systems beruht.[521]
Dies ergibt sich insbesondere aus dem Rechtsstaatsprinzip gemäß Art. 20 Abs. 3 GG unter dem Gesichtspunkt der Transparenz staatlichen Handelns,[522] dem Anspruch auf ein faires Verfahren gemäß Art. 2 Abs. 1, 20 Abs. 3 GG, aus Art. 6 Abs. 1 EMRK,[523] Art. 47 Abs. 1, Abs. 2 GRCh[524] und aus Art. 8, 15 Abs. 2 KI-Konvention.[525] Darüber hinaus sind insbesondere die Regelungen und Pflichten nach Art. 50 ff. KI-VO zu beachten.[526] Bei Einordnung eines Unterstützungssystems nach Funktionalität in den Hochrisiko-Bereich bzw. bei entscheidungsersetzenden Systemen sind insbesondere die Pflichten für Betreiber solcher Systeme nach Art. 26 KI-VO bzw. für Anbieter solcher Systeme nach Art. 13 KI-VO und Art. 14 Abs. 4 KI-VO zu berücksichtigen, die weitreichender sein können als jene nach Art. 50 ff. KI-VO bzw. jene nach geltendem nationalem Recht. Auch ist bei Einsatz eines Hochrisiko-KI-Systems eine Grundrechtefolgenabschätzung nach Art. 27 KI-VO durchzuführen, die an die nationale Marktüberwachungsbehörde i. S. d. Art. 70 Abs. 1 KI-VO zu übermitteln ist.
Eine Pflicht zur Information über den verwendeten Quellcode besteht nach unseren Erkenntnissen hingegen grundsätzlich nicht. Einer solchen Pflicht stehen regelmäßig die Geheimhaltungsinteressen des Entwicklers oder der Entwicklerin bzw.

[518] Siehe hierzu ausführlich *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025.
[519] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 99f., 142ff., (s. a. insb. 154f.), 258ff., 265ff., 321ff.
[520] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 146ff., 258ff., 321ff.
[521] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 146ff., 258ff., 321ff., 355ff.
[522] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 146ff. (s. a. insb. 154f.).
[523] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 241ff., 355ff.
[524] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 265ff.
[525] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 258ff.
[526] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 289ff.

Sicherheitsinteressen entgegen, wobei Ausnahmekonstellationen denkbar sind, in denen eine Offenlegung relevant werden kann.[527]
Dass Systeme, die auf maschinellem Lernen beruhen, nicht über eine vollumfänglich erklärbare und nachvollziehbare innere Logik verfügen, sog. Black-Box-Problematik, steht der Erfüllung einer Transparenz- und Informationspflicht nicht entgegen, da insoweit eine nachträgliche Entscheidungsbegründung dem Transparenzgebot hinreichend Rechnung trägt.[528] Regelbasierte Systeme weisen dem Grunde nach eine bessere Nachvollziehbarkeit auf als Systeme maschinellen Lernens.[529] Gleichwohl können jene auch mit Blick auf eine detailscharfe Erklärbarkeit aller Vorgänge schnell hochkomplex werden[530], indes erscheinen sie vor dem Hintergrund von Transparenz und Information als vorteilhaft(er gegenüber Systemen maschinellen Lernens). Auch ihre abstrakte Funktionalität ist allerdings zur Kenntnis der Beteiligten bzw. der Öffentlichkeit zu bringen, auch wenn nicht der vollständige Entscheidungsbaum/Quellcode offenzulegen ist.

b. Handlungsoptionen und -empfehlungen

Die Handlungsoptionen und Handlungsempfehlungen bezüglich der Transparenz beim Einsatz eines KI-Systems bei der richterlichen Entscheidungsfindung lassen sich nach den Adressaten der zu erteilenden Informationen ordnen. Unterschieden werden können hierbei zum einen die sog. **äußere Transparenz** gegenüber den Parteien und der Öffentlichkeit und zum anderen die sog. **innere Transparenz** gegenüber den das System nutzenden Richtern und Richterinnen sowie dem Betreiber. Die äußere Transparenz wird zuerst behandelt und ergibt sich aus verfassungsrechtlichen Vorgaben sowie insbesondere aus Art. 26 Abs. 11 KI-VO, Art. 50 ff. KI-VO und Art. 86 KI-VO. Sodann werden Optionen und Empfehlungen zur inneren Transparenz gegenüber den das System nutzenden Richtern und Richterinnen sowie dem Betreiber, insbesondere im Hinblick auf die sog. Black-Box-Problematik und die Anforderungen des Art. 13 KI-VO, erörtert. Abschließend werden die Instrumente zur Überprüfbarkeit des Systems und zur Dokumentation wie z. B. eine Quellcode-Offenlegung und Protokollierung dargestellt.

aa. Äußere Transparenz gegenüber Parteien und Öffentlichkeit

In dieser Lesart sind Adressaten der äußeren Transparenz die Parteien des Verfahrens sowie die Öffentlichkeit. Neben der aus dem nationalen Recht folgenden Information über ein „Ob“ und ein „Wie“ des Einsatzes sind die Transparenzpflichten der Art. 50 ff. KI-VO zu beachten.

---

[527] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 99ff., 146ff., (s. a. insb. 154f.), 258ff., 265ff., 321f., 351, 355ff.; siehe auch *Gsell/Pfündl*, RDi 2026, 457, 463, die de lege ferenda eine Pflicht befürworten, dass Parteien durch das Gericht auf Art und Umfang des KI-Einsatzes hinweisen, unter Hinweis, dass Richter die Parteien bereits jetzt gem. § 139 Abs. 2 ZPO auf den Einsatz eines KI-Tools bei der Entscheidungsfindung, -erstellung bzw. -redaktion hinweisen müsse.
[528] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 118f., 148ff., 262, 351.
[529] Vgl. *Wischmeyer*, AöR 143 (2018), 1, 46f.
[530] Vgl. *Wischmeyer*, AöR 143 (2018), 1, 46.

(1) Information über das „Ob“ und das „Wie“ des Einsatzes

Die Parteien und die Öffentlichkeit sind über das „Ob“ und das „Wie“ des Einsatzes von informationstechnischen Systemen im Prozess zu informieren.[531] Mit Blick auf das „Ob“ des Einsatzes **sollte eine Mitteilung vor bzw. spätestens zum Verfahrensbeginn** erfolgen. In der Literatur wird auch eine Ablehnungsmöglichkeit des Einsatzes vorgeschlagen, was allerdings nicht zwingend ist.[532] Auch eine regelmäßige Wiederholung dieser Mitteilung, wie sie in der Literatur vorgeschlagen wird,[533] ist nicht zwingend erforderlich, kann allerdings das Bewusstsein für die Kontrollnotwendigkeiten erhöhen. Umfassen muss eine solche Mitteilung allerdings die Information, welcher Art das System ist, d. h. ob regelbasierte und/oder Systeme maschinellen Lernens zum Einsatz kommen[534] und welche abstrakte Funktionslogik diesem System unterliegt.
Bei der Ausgestaltung der Information der Öffentlichkeit und der Parteien sollten auch die Maßnahmen anderer Rechtskreise im Blick behalten und ggf. adaptiert werden. So haben in Frankreich Behörden die Pflicht, die Funktionsweise und wesentliche Eigenschaften algorithmischer Unterstützung offenzulegen, was das Maß der Einflussnahme der algorithmischen Verfahren auf die behördliche Entscheidung, die verarbeiteten Daten und Datenquellen, Parameter und Gewichtung sowie durchgeführten Aufgaben des Systems erfasst.[535]
Derartige Maßnahmen können bzw. sollten auch hierzulande im judikativen Kontext adaptiert werden. Mit Blick auf die **genaue Ausgestaltung der Information** von Parteien und Öffentlichkeit sollte beachtet werden, dass **keine „Informationsüberflutung“** erfolgt, die dazu führen kann, dass eine „Desinformation durch Information“ entsteht und diese Gruppen die ihnen mitgeteilten Informationen auch tatsächlich wahrnehmen (können).[536] Im Rahmen der Informationsverarbeitung sind zeitliche Restriktionen und konkurrierende Bedürfnisse zu beachten; zu viele Informationen können erhebliche Transaktionskosten bei den Adressaten hervorrufen und den Blick auf das Wesentliche verstellen sowie Abwehrreaktionen hervorrufen.[537] Hierbei sollte beachtet werden, dass feine juristische sowie technische Formulierungen für die allgemeine Verständlichkeit und Informationsaufnahme nicht immer vorteilhaft sein können. Es sollte daher auf verständliche Sprache gesetzt werden.[538]
Eine Handlungsoption besteht vor diesem Hintergrund in einem **gestuften Informationsmodell**: Die Informationen sollten gezielt für verschiedene Gruppen und in verschiedenen Abstufungen an inhaltlicher Tiefe sowie sprachlicher Schwierigkeit

[531] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 147f.
[532] *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 338f.
[533] Vgl. im behördlichen Kontext *Martini*, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz, 1. Aufl. 2019, S. 180.
[534] Vgl. im behördlichen Kontext *Martini*, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz, 1. Aufl. 2019, S. 180.
[535] *Martini*, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz, 1. Aufl. 2019, S. 186.
[536] *Martini*, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz, 1. Aufl. 2019, S. 188.
[537] *Martini*, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz, 1. Aufl. 2019, S.188.
[538] *Martini*, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz, 1. Aufl. 2019, S. 188f.

zur Verfügung gestellt werden.[539] Es sollte beachtet werden, dass mit steigender System- und Einsatzkomplexität grundsätzlich höhere Anforderungen an die Information einhergehen.[540]
Die notwendigen Informationen (wie z. B. zur Erklärung zur abstrakten Funktionalität) können **barrierefrei auf (den) Webseiten (der Justiz), in Textform (etwa Leitfäden) sowie in graphisch oder per Video aufbereiteter Form bereitgestellt werden**.[541] Bei der bildlichen Aufbereitung sollte insbesondere auf intuitiv erfassbare Bildsymbolik mit hinreichender Größe und Aussagekraft geachtet werden.[542] Die Videos können insbesondere die Funktionen erklären und zeigen, wie die Anwendungen und Abläufe (im Groben) in der Praxis aussehen, wobei angesichts der Akzeptanz insgesamt hierbei ein großer Wert darauf gelegt werden sollte, zu zeigen, dass gerade im entscheidungsunterstützenden Einsatzbereich die menschlichen Richterinnen und Richter weiterhin die Hoheit bzw. einen/den wesentlichen Einfluss auf den Prozess der Entscheidungsfindung sowie auf die Entscheidung selbst haben. Entsprechendes gilt für den entscheidungsersetzenden Einsatz mit Blick auf die Darstellung des Ausmaßes der menschlichen Aufsicht und Kontrolle bei der Überwachung des automatisierten Entscheidungsvorgangs. Mit Blick auf die Barrierefreiheit des Rechtsschutzes und auf möglicherweise fehlende digitale Kompetenzen[543], sollten die Informationen auch physisch bereitgestellt werden (können), **beispielsweise mittels postalischer Übersendung** bzw. physischer Hinterlegung bei den KI einsetzenden staatlichen Institutionen.
Informationen können im Verfahrenskontext zusätzlich **in gerichtlichen Schreiben**, wie beispielsweise richterlichen Hinweisen oder Ladungen aufgenommen werden. Auch können diese direkt mit den auf der Website abrufbaren Informationen verlinkt werden, was die Zugänglichkeit niederschwellig gestaltet. Auch bietet es sich an, standardisierte Hinweise bereits zu Verfahrensbeginn sowie im Falle des Hinzutretens weiterer Beteiligter erneut zu übermitteln sowie bei einer wesentlichen Änderung des Einsatzszenarios diese Hinweise aktualisiert bereitzustellen.[544] Auch wäre es möglich, die notwendigen Informationen auf einem **bundeseinheitlichen Justizportal**[545] zum

---

[539] Vgl. *Wischmeyer*, AöR 143 (2018), 1, 46ff., 63; vgl. *Dörr*, Elektronischer Rechtsverkehr und Bürgerzugänge, in: Reuß/Windau (Hrsg.), Göttinger Kolloquien zur Digitalisierung des Zivilverfahrensrechts, Kolloquien im Sommersemester 2021, Göttingen 2022, S. 118; vgl. zum behördlichen Kontext *Martini*, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz, 1. Aufl. 2019, S. 180.
[540] Chibanguza/Kuß/Steege KI-Hdb/*Jungkind/S. Koch*, 2. Aufl. 2026, 2. Teil: § 4 D. Rn. 62.
[541] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 147; zur Barrierefreiheit *Bundesministerium für Arbeit und Soziales/Bundesfachstelle Barrierefreiheit*, Leitfaden „Umsetzung digitaler Barrierefreiheit“, Februar 2026, abrufbar unter https://www.esf-bw.de/esf-plus-in-bw/querschnittsziele-und-grundlegende-voraussetzungen/chancengleichheit-und-nichtdiskriminierung (Abruf am 23.09.2026, 21:22 Uhr), S. 8ff.; vgl. auch *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 6.
[542] *Martini*, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz, 1. Aufl. 2019, S. 189.
[543] Siehe etwa mit Blick auf elektronische Kommunikationswege und darauf, dass in Deutschland lebende Personen nicht grundsätzlich an digitale Verfahrensweisen gewöhnt sind, Ebers/Heinze/Steinrötter KI/*Huber/Giesecke*, 2. Aufl. 2026, § 9 Rn. 8.
[544] Vgl. Art. 50 Abs. 5 KI-VO, wonach Informationen spätestens zum Zeitpunkt der ersten Interaktion oder Aussetzung in klarer und eindeutiger Weise barrierefrei bereitzustellen sind.
[545] Hierzu vgl. die *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, Münchener Thesen zum Zivilprozess der Zukunft, abrufbar unter https://www.olg-duesseldorf.nrw.de/infos/Zivilprozess-der-Zukunft/Muenchener-Thesen-zum-Zivilprozess-der-Zukunft.pdf (Abruf am 26.02.2026, 18:27 Uhr), S. 6, These A.3; siehe auch §§ 1131ff. ZPO, insbesondere § 1131 Abs. 3 S. 1 ZPO.

Abruf bereitzustellen, sodass die Informationen an einem Ort zentral vorgehalten werden können.
Soweit unter Anwendung eines entscheidungs*ersetzenden* informationstechnischen Systems Parteien ausschließlich mit dem System bzw. Online-Portalen interagieren, könnte als weitere Handlungsoption eine zwingende Bestätigung der Vermittlung der zu verschaffenden Informationen vorgesehen werden, bevor das System den Fortschritt des Verfahrens freigibt. Die Kenntnisnahme der zu vermittelnden Informationen müsste somit vor Verfahrensdurchführung durch die Partei bestätigt werden. Dies würde zumindest bei der klagenden Partei die Kenntnisnahme(möglichkeit) sicherstellen. Hinsichtlich der beklagten Partei müsste eine solche zwingende Bestätigung freilich an die Geltendmachung der Verteidigungsmittel geknüpft sein, denn andernfalls würde ihr in Anbetracht der Klage ein unberechtigtes Blockadepotential zukommen. Die beklagte Partei könnte durch Nichtbestätigung den Fortgang des Verfahrens hindern.
Die elektronische Akte ist in der Justiz grundsätzlich verpflichtend, siehe etwa § 298a ZPO, seit dem 01.01.2026; die beA-Nutzungspflicht für Rechtsanwältinnen und Rechtsanwälte, siehe etwa § 130d ZPO, besteht seit dem 01.01.2022. Die Justiz insgesamt, insbesondere auch die Rechtsanwältinnen und Rechtsanwälte werden zunehmend in den Digitalisierungsprozess eingebunden. Als Handlungsoption bietet es sich daher flankierend an, dass Rechtsanwältinnen und Rechtsanwälte ermutigt sowie angeregt werden, in ihrer Beratung ihre Mandantschaft über das „Ob“ des Einsatzes von informationstechnischen Systemen bei der richterlichen Entscheidungsfindung zu informieren und auf die entsprechend justizseitig vorzuhaltenden Informationen hinzuweisen. Rechtsanwältinnen und Rechtsanwälte könnten als vertrauenswürdige Intermediäre zwischen Rechtssphäre und Mandantschaft (somit) die Akzeptanz des Einsatzes von KI in der Justiz unterstützen.
Eine Pflicht zur Information der Parteien folgt auch aus Art. 26 Abs. 11 KI-VO. Nach dieser Vorschrift sind natürliche Personen als Betroffene von Entscheidungen von Hochrisiko-KI-Systemen oder beim unterstützenden Einsatz der Systeme bei Entscheidungen über sie vom Betreiber zu informieren. Dementsprechend muss die Ziviljustiz als Betreiber[546] darüber informieren, dass sie ein konkretes System bei der richterlichen Entscheidungsfindung einsetzt.[547] Macht eine der Parteien ihr Recht auf Erklärung nach Art. 86 KI-VO geltend, muss der Betreiber eine klare und aussagekräftige Erläuterung zur Rolle des KI-Systems im Entscheidungsprozess und zu den wichtigsten Elementen der getroffenen Entscheidung abgeben.[548] Eine solche Erläuterung sollte vorgehalten und fortlaufend aktualisiert werden. Weitergehende als die oben beschriebenen Pflichten folgen somit aus Art. 26 Abs. 11 KI-VO nicht.
Bereits vor dem finalen Einsatz kann die Öffentlichkeit bzw. können die Parteien zu informieren sein. Dies ist der Fall, wenn ein Hochrisiko-KI-System außerhalb eines KI-Reallabors getestet wird. Die Testteilnehmer müssen dann nach Art. 60 Abs. 4 lit. i KI-VO ihre informierte Einwilligung erteilen. Die informierte Einwilligung ist im Detail in Art. 61 KI-VO geregelt. Die Testteilnehmer müssen mit präzisen, klaren, relevanten und verständlichen Informationen über die Inhalte der Art. 61 Abs. 1 lit. a–e KI-VO informiert werden. Hinzu tritt eine Pflicht zur Datierung und Dokumentation der informierten Einwilligung. Den Testteilnehmern oder deren gesetzlichen Vertretern ist eine Kopie auszuhändigen, vgl. Art. 61 Abs. 2 KI-VO.

---

[546] Zur Einordnung: *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 280; ebenso *Wachter/Leeb*, RDi 2024, 440, 441.
[547] Martini/Wendehorst/*I. Eisenberger/Th. Binder*, 2. Aufl. 2026, KI-VO Art. 26 Rn. 104.
[548] Martini/Wendehorst/*I. Eisenberger/Th. Binder*, 2. Aufl. 2026, KI-VO Art. 26 Rn. 104.

Informationen können freilich erst gewährt werden, sobald diese vorliegen, so kann beispielsweise vor der Verarbeitung nicht über das (zukünftige) Ergebnis informiert werden.[549] Im Interesse einer größtmöglichen Informiertheit sollten Informationen jedoch so früh und so umfassend wie möglich zur Verfügung gestellt werden. Im Sinne einer zeitlichen Stufung können an geeigneten Stellen im Verfahren jeweils konkretere sowie detailliertere Informationen zur Verfügung gestellt werden, sobald diese vorliegen.[550]

(2) Transparenzpflichten nach Art. 50 ff. KI-VO

Transparenzpflichten ergeben sich ferner aus den Art. 50 ff. KI-VO. Viele (Transparenz)Pflichten der KI-VO gelten im Hochrisiko-Bereich gem. Art. 113 KI-VO in der Fassung der Verordnung (EU) 2026/1744 („Digital Omnibus“) erst ab dem 02.12.2027 (Hochrisiko-KI-Systeme nach Anhang III) bzw. ab dem 02.08.2028 (Anhang I). [551] Die KI-VO unterteilt KI-Systeme in Risiko-Kategorien, von der höchsten Kategorie, dem Verbot von Formen des KI-Einsatzes nach Art. 5 KI-VO, über Hochrisiko-KI-Systeme nach Art. 6 KI-VO, die KI-Systeme mit begrenztem Risiko bis hin zur niedrigsten Risiko-Stufe, den KI-Anwendungen mit minimalen Risiken, bei denen auf rechtliche Anforderungen verzichtet wird. Mit Blick auf den Aspekt der Transparenz und Information sind insbesondere die **Transparenzpflichten nach Art. 50 KI-VO**, der erst seit dem 02.08.2026 gilt, in den Blick zu nehmen, deren Vorgaben ab der zweitniedrigsten Risiko-Kategorie, den KI-Systemen mit begrenztem Risiko und damit auch im Falle der Einordnung als Hochrisiko-KI-System und somit neben den daraus folgenden (weiteren) Pflichten, zu befolgen sind, s. a. Art. 50 Abs. 6 KI-VO.[552] Während Art. 50 Abs. 1 und 2 KI-VO Anbieter adressieren, befassen sich Art. 50 Abs. 3 und 4 KI-VO mit dem Betreiber. Beide Komplexe sind im Rahmen informationstechnischer Systeme für die richterliche Entscheidungsfindung relevant.[553] Art. 50 Abs. 1 bis 4 KI-VO haben primär das „Ob“ der Kennzeichnung KI-generierten Inhalts bzw. der Interaktion mit einem KI-System zum Gegenstand. Hiernach haben gem. Art. 50 **Abs. 1** KI-VO **Anbieter** von KI-Systemen, die zur **direkten Interaktion mit natürlichen Personen** bestimmt sind, dafür Sorge zu tragen, dass diese darüber **informiert** werden, dass sie mit einem **KI-System interagieren**.[554] Art. 50 **Abs. 2** KI-VO schreibt **Anbietern** von KI-Systemen (auch mit allgemeinem Verwendungszweck) vor, dass die **Systemausgaben maschinenlesbar gekennzeichnet und als künstlich erzeugt bzw. manipuliert erkennbar** sind.[555] Betreiber von Emotionserkennungssystemen und von KI-Systemen zur biometrischen

[549] Paal/Pauly/*Hennemann*, 4. Aufl. 2026, VO (EU) 2016/679 Art. 13 Rn. 31a.
[550] Vgl. Ebers/Heinze/Steinrötter KI/*Steinrötter/Pfeiffenbring*, 2. Aufl. 2026, § 18 Rn. 88f.
[551] Hingewiesen sei an dieser Stelle auch auf die Regelungen des Art. 111 KI-VO.
[552] BeckOK KI-Recht/*Lauber-Rönsberg*, 7. Ed. 25.7.2026, KI-VO Art. 50 Rn. 2.
[553] Hilgendorf/Roth-Isigkeit KI-VO/*Kumkar*, 2. Aufl. 2025, § 6 Rn. 37; *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 278ff.; vgl. auch Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 22; vgl. auch *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 10.
[554] Bomhard/Pieper/Wende/*Merkle*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 23.
[555] Bomhard/Pieper/Wende/*Merkle*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 23.

Kategorisierung müssen nach Art. 50 Abs. 3 KI-VO die betroffenen natürlichen Personen über den Betrieb informieren und dürfen personenbezogene Daten nur datenschutzrechtskonform verarbeiten.[556] Art. 50 **Abs. 4 UAbs. 1** KI-VO enthält eine **Offenlegungspflicht** für die **Betreiber von Deepfake**-Bild-, Ton- oder Videoerzeugungs- oder -manipulations-KI-Systemen hinsichtlich der **künstlichen Erzeugung oder Manipulation** der Inhalte.[557] Art. 50 **Abs. 4 UAbs. 2** KI-VO verpflichtet **Betreiber** von KI-Systemen, die **Text erzeugen oder manipulieren**, der **veröffentlicht** wird, um die **Öffentlichkeit** über **Angelegenheiten von öffentlichem Interesse zu informieren**, dazu, diese künstliche **Erzeugung oder Manipulation offenzulegen**.[558] Das „Wie“ der Kennzeichnung wird hingegen maßgeblich in Art. 50 Abs. 5 KI-VO bestimmt, wonach die Informationen den betreffenden natürlichen Personen **spätestens zum Zeitpunkt der ersten Interaktion oder Aussetzung** in klarer und eindeutiger Weise bereitgestellt werden müssen sowie den geltenden Barrierefreiheitsanforderungen entsprechen müssen.[559]
Als **zu empfehlende Handlungsoption** kommt in Betracht, die notwendigen **Transparenzpflichten in die prozessuale Verfahrenskommunikation bzw. Arbeitskommunikation direkt einzubetten** und somit separate Einzelhinweise zu vermeiden, die potentiell verwirrungsstiftend sein können.

(i) Offenlegung der Interaktion mit einem KI-System (Anbieterpflicht)

Nach Art. 50 Abs. 1 KI-VO müssen Anbieter sicherstellen, dass KI-Systeme, die für die direkte Interaktion mit natürlichen Personen bestimmt sind, so konzipiert und entwickelt werden, dass die betreffenden **natürlichen Personen informiert** werden, dass sie **mit einem KI-System interagieren**, es sei denn, dies ist aus Sicht einer angemessen informierten, aufmerksamen und verständigen natürlichen Person aufgrund der Umstände und des Kontexts der Nutzung offensichtlich. Ziel der Regelung ist insbesondere, den Risiken einer Täuschung zu begegnen, die aus der Interaktion mit einem informationstechnischen System resultieren können.[560] Natürlichen Personen muss also klargemacht werden, dass sie mit einem KI-System, nicht mit einem Menschen interagieren.
Die KI-Systeme, die unter diese Verpflichtung fallen, müssen für die direkte Interaktion mit natürlichen Personen bestimmt sein; diese Bestimmung liegt nach einer Auffassung in der Literatur dann nicht vor, wenn die KI-System-Ergebnisse vor der Interaktion mit der betreffenden Person durch einen anderen Menschen validiert bzw. weiterverarbeitet wurden.[561] Das „Bestimmen“ bemisst sich grundsätzlich an der subjektiven Vorgabe des Anbieters hinsichtlich der Einsatzszenarien sowie der Einsatzweise, begrenzt durch eine objektive Willkürkontrolle.[562] Die Norm erfasst insbesondere Chat-Bots, ist jedoch technologieoffen formuliert.[563] Die Offenlegung sollte daher insbesondere bei Dialogsystemen erfolgen[564], aber auch im Übrigen sollte die Interaktion mit einem – andere Funktionen beinhaltenden – KI-System deutlich

[556] Bomhard/Pieper/Wende/*Merkle*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 24.
[557] Bomhard/Pieper/Wende/*Merkle*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 24.
[558] Bomhard/Pieper/Wende/*Merkle*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 24.
[559] *Deutscher Bundestag, Fachbereich Europa*, Zu rechtlichen Vorgaben der Verordnung über Künstliche Intelligenz für den Einsatz von KI in Behörden der EU-Mitgliedstaaten, Infobrief EU 6 - 3010 - 032/25, 07.07.2025, S. 38f.
[560] BeckOK KI-Recht/*Lauber-Rönsberg*, 7. Ed. 25.7.2026, KI-VO Art. 50 Rn. 15.
[561] BeckOK KI-Recht/*Lauber-Rönsberg*, 7. Ed. 25.7.2026, KI-VO Art. 50 Rn. 12ff.
[562] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 50 Rn. 59.
[563] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 50 Rn. 57.
[564] Vgl. BeckOK KI-Recht/*Lauber-Rönsberg*, 7. Ed. 25.7.2026, KI-VO Art. 50 Rn. 14.

gemacht werden. Insbesondere mit Blick auf den Einsatz informationstechnischer Systeme beim entscheidungs*ersetzenden* Einsatz ist zusätzlich zu beachten, dass bei einer Interaktion der Systeme mit schutzbedürftigen Personengruppen, deren Merkmale berücksichtigt werden sollen, Erwägungsgrund 132 S. 4 KI-VO. Explizit aufgeführt sind hierbei die Angehörigkeit zu schutzbedürftigen Gruppen aufgrund von Alter oder Behinderung. Dies bedeutet konkret, dass die Offenlegung den geltenden Barrierefreiheitsanforderungen[565] – Art. 50 KI-VO selbst regelt keine eigenständigen oder zusätzlichen derartigen Anforderungen – entspricht und etwa behinderte Personen abgestimmt auf ihre Behinderungen die Informationen in einem zugänglichen Format erhalten (zur Barrierefreiheit des Rechtsschutzes siehe zudem Abschnitt C.II.8).[566]

Im Sinne einer Implementierung in der prozessualen Verfahrenskommunikation kann daher ein **steter Hinweis mit Bestätigung der Kenntnisnahme** hiervon, vor jeder Interaktion mit dem entscheidungs*ersetzenden* System implementiert werden, dass die Partei mit einer KI interagiert. Dieser Hinweis kann zudem **prominent in die gerichtsseitige Korrespondenz** aufgenommen werden und **wiederholt bei einer Interaktion mit dem System angezeigt bzw. ausgegeben** werden. Mit Blick auf vulnerable Personengruppen kann auch eine partielle Anwaltspflicht für Prozesse, die von entscheidungsersetzenden informationstechnischen Systemen geführt werden, angedacht werden, solange noch keine hinreichenden Möglichkeiten zur gerichtsseitigen Adressierung der individuellen Schutzbedürftigkeit geschaffen sind, damit die individuellen Belange und Interessen durch den anwaltlichen Beistand gewahrt werden können. In diesen Fällen sollte angedacht werden, die zusätzlichen Kosten, die durch den pflichtigen anwaltlichen Beistand (nur) bei den entsprechenden schutzbedürftigen Personengruppen entstehen, durch andere Maßnahmen zu kompensieren, z. B. durch Gewährung von Prozesskostenhilfe. Diese Handlungsoption sollte jedoch primär als Übergangslösung betrachtet und zügig Abhilfe geschaffen werden.

Nach Art. 50 Abs. 1 S. 1 Hs. 2 KI-VO besteht eine **Ausnahme von der Pflicht zur Offenlegung der Interaktion mit einem KI-System**, sofern dies aus Sicht einer angemessen informierten, aufmerksamen und verständigen natürlichen Person aufgrund der Umstände und des Kontexts der Nutzung offensichtlich ist. Aufgrund der Wendung des Wortlauts „es sei denn“ kann mit der Verortung der Darlegungs- und Beweislast für die Offensichtlichkeit beim Anbieter gerechnet werden.[567] Die Bestimmung der entsprechend qualifizierten natürlichen Person orientiert sich am europäischen Verbraucherleitbild, dem durchschnittlich informierten, aufmerksamen

[565] Siehe hierzu u. a. Richtlinie (EU) 2016/2102 über den barrierefreien Zugang zu den Websites und mobilen Anwendungen öffentlicher Stellen sowie die Richtlinie (EU) 2019/882 des Europäischen Parlaments und des Rates über die Barrierefreiheitsanforderungen für Produkte und Dienstleistungen; siehe auch *Europäische Kommission,* Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme nach Artikel 50 der KI-Verordnung der Europäischen Kommission „Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the ‘AI Act’)“, abrufbar unter https://digital-strategy.ec.europa.eu/de/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems (Abruf am 31.08.2026, 20:51 Uhr), Rn. 144.

[566] So auch *Europäische Kommission,* Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme nach Artikel 50 der KI-Verordnung der Europäischen Kommission „Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the ‘AI Act’)“, abrufbar unter https://digital-strategy.ec.europa.eu/de/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems (Abruf am 31.08.2026, 20:51 Uhr), Rn. 34, 144.

[567] BeckOK KI-Recht/*Lauber-Rönsberg*, 7. Ed. 25.7.2026, KI-VO Art. 50 Rn. 19.

sowie verständigen Durchschnittsverbraucher.[568] Hierzu sollte in concreto die Zielgruppe betrachtet werden, mit der das KI-System interagiert (bzw. das [ggf. breitere] Publikum, das bestimmungsgemäß und vernünftigerweise wahrscheinlich mit diesem interagieren wird) und sodann das durchschnittliche Mitglied als Maßstab genommen werden.[569] Zur Beurteilung der Offensichtlichkeit sind gem. Art. 50 Abs. 1 S. 1 KI-VO die Umstände und der Nutzungskontext relevant. Als relevante Faktoren können hierbei insbesondere herangezogen werden, die Systembezeichnung, das Erscheinungsbild des Systems, seine Reaktionsgeschwindigkeit sowie die Systemfunktionen und der Einsatzbereich des Systems.[570]

Im entscheidungs*ersetzenden* Einsatzszenario ist durchaus vorstellbar, dass rechtssuchende Bürgerinnen und Bürger, die mit einem online auf einem Justizportal bereitgestellten automatisierten Entscheidungssystem agieren, aus dem Kontext erkennen können, dass sie mit einem KI-System interagieren. Selbstverständlich ist dies jedoch nicht, insbesondere dann nicht, wenn die Interaktion über Chat-Bots gestaltet ist. Da Chat-Interaktionen auch über natürliche Personen abgewickelt werden können, wird sich die Offensichtlichkeit des KI-Einsatzes dann nicht einstellen. Ein Hinweis ist somit in diesen Konstellationen erforderlich.[571] Mit Blick auf den bei schutzbedürftigen Gruppen erhöhten Maßstab sowie dem vorherrschenden, alleinigen Verständnis richterlicher Tätigkeit als menschlicher Tätigkeit, **ist zu empfehlen,** im entscheidungs*ersetzenden* Einsatz informationstechnischer Systeme **stets die konkrete Offenlegung der KI-Interaktion** in klarer Weise, d. h. in allgemein verständlicher, und eindeutiger, d. h. das Risiko einer Verwechslung ausschließender Form (z. B. mittels einer gut sichtbaren Kopfzeile oder einer dem KI-System vorgeschalteten Eingangssequenz), vorzusehen (siehe Art. 50 Abs. 5 S. 1 KI-VO).[572]

Anders zu bewerten ist dies allerdings bei einem entscheidungs*unterstützenden* Einsatz solcher Systeme. Bei einem solchen Einsatzszenario wäre die KI-Interaktion gegenüber den dieses System nutzenden Richterinnen und Richtern offenzulegen, da insoweit eine Interaktion vorgenommen wird. Keine Offenlegungspflicht besteht demgegenüber im Verhältnis zu den Parteien, da diese nicht direkt mit dem KI-System interagieren.[573] Gegenüber nutzenden Richterinnen und Richtern erscheint ein gesonderter Hinweis auf die Interaktion mit einem KI-System allerdings regelmäßig verzichtbar, da sich die KI-Interaktion regelmäßig aus dem Nutzungskontext offensichtlich ergibt. Wenn Richterinnen und Richter bewusst ein KI-Unterstützungssystem wie die MAKI-KI zur Unterstützung ihrer Entscheidungsfindung

---

[568] Bomhard/Pieper/Wende/*Merkle*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 48f.; BeckOK KI-Recht/*Lauber-Rönsberg*, 7. Ed. 25.7.2026, KI-VO Art. 50 Rn. 20.

[569] Vgl. *Europäische Kommission,* Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme nach Artikel 50 der KI-Verordnung „Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the ‘AI Act’)“, abrufbar unter https://digital-strategy.ec.europa.eu/de/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems (Abruf am 31.08.2026, 20:51 Uhr), Rn. 44; vgl. auch Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 50 Rn. 63.

[570] Hierzu ausführlich Bomhard/Pieper/Wende/*Merkle*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 51ff.

[571] Vgl. auch *Europäische Kommission,* Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme nach Artikel 50 der KI-Verordnung „Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the ‘AI Act’)“, abrufbar unter https://digital-strategy.ec.europa.eu/de/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems (Abruf am 31.08.2026, 20:51 Uhr), Rn. 45, die einen (sehr) restriktiven Ansatz verfolgen.

[572] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 50 Rn. 155ff.

[573] BeckOK KI-Recht/*Lauber-Rönsberg*, 7. Ed. 25.7.2026, KI-VO Art. 50 Rn. 13f. mit einem Beispiel anhand eines Schachcomputers.

in einem konkreten Rechtsstreit nutzen, sind sie allein schon aufgrund der Bezeichnung des Systems als KI-Unterstützungssystem über die Tatsache der Interaktion mit dem KI-System informiert, sodass sich die Interaktion aus dem Nutzungskontext unweigerlich ergibt. Ein gesonderter Hinweis ist diesen Richterinnen und Richtern gegenüber damit obsolet. Dem/der durchschnittlichen Richter/in, der/die mit dem informationstechnischen System im richterlichen Kontext interagiert, dürfte die Interaktion mit einer KI offensichtlich sein. Mit Blick auf den Zweck der Regelung, dem besonderen Risiko in Bezug auf Identitätsbetrug bzw. Täuschung bestimmter KI-Systeme durch Transparenz zu begegnen, Erwägungsgrund 132 S. 1 u. 2 der KI-VO, dürfte das Risiko von nur geringer Bedeutung sein und dort allenfalls im Dialogsystem-Einsatz.[574] Gleichwohl spricht auch hier nichts gegen eine freiwillige adäquate Information (z. B. über Hinweisbanner) über die KI-Interaktion mittels geeigneter Hinweise. Entsprechendes gilt bei unterstützenden Systemen, die agentische Verfahrensassistenten vorsehen. Auch hier ist in der Regel eine KI-Interaktion offensichtlich, ein nochmals konkreter Hinweis auf die KI-Interaktion schadet auch hier allerdings nicht.

(ii) Kennzeichnungspflicht „synthetisch“ generierter Textinhalte (Anbieterpflicht)

Art. 50 Abs. 2 KI-VO[575] legt darüber hinaus eine weitere Pflicht fest, indem die Norm Anbieter von KI-Systemen dazu verpflichtet, „synthetisch“ generierte Textinhalte als solche zu **kennzeichnen.** D. h., wenn das KI-System **„künstlich hergestellte“ (also nicht-menschlich erzeugte) Inhalte** ausgibt,[576] sind diese Inhalte als solche zu bezeichnen und kenntlich zu machen. Mit der Regelung soll der fehlenden Unterscheidbarkeit von Inhalten begegnet werden, die von Menschen natürlich bzw. von Maschinen künstlich generiert werden, sowie die Integrität des gesamten Informationsökosystems und das ihm entgegengebrachte Vertrauen geschützt werden.[577] Die Kennzeichnung soll daher gem. Art. 50 Abs. 2 S. 1 a. E. KI-VO nicht nur für Menschen, sondern gerade wegen der weiteren Interaktionsmöglichkeiten mit weiteren KI-Systemen, auch für Maschinen lesbar sein,[578] was letztlich die Authentizität der Systemergebnisse auch gegenüber weiterverarbeitenden Systemen sichern soll.
Die Einzelheiten und Maßstäbe für die Wirksamkeit, Interoperabilität, Belastbarkeit und Zuverlässigkeit der technischen Lösungen hierfür sind in Art. 50 Abs. 2 S. 2 KI-VO

---

[574] Vgl. Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 50 Rn. 63.
[575] Ausführlich hierzu *Molavi Vasse’i*, RDi 2024, 406, 406ff.
[576] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 289; Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 50 Rn. 73.
[577] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 289; Erwägungsgrund 133 S. 2 der KI-VO.
[578] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 289; sowie BeckOK KI-Recht/*Lauber-Rönsberg*, 7. Ed. 25.7.2026, KI-VO Art. 50 Rn. 34 hierfür zusätzlich den Wortlaut von Erwägungsgrund 133 S. 3 der KI-VO anführend; nur die maschinelle Lesbarkeit und nicht eine menschliche nennend HK-KI-VO/Hilgendorf/Härtlein/*Hilgendorf/Härtlein*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 10; der nur maschinellen Lesbarkeit zuneigend, aber auch sich der menschlichen gegenüber nicht verschließend *Europäische Kommission,* Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme nach Artikel 50 der KI-Verordnung „Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the ‘AI Act’)“, abrufbar unter https://digital-strategy.ec.europa.eu/de/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems (Abruf am 31.08.2026, 20:51 Uhr), Rn. 71 („Perceptible marks and labels are not excluded as a complementary measure […]“), 77 („This information is **notably** [= insbesondere; Hervorhebung durch den Bearbeiter] the result of the detection […]“).

aufgeführt, der Anbietern eine Erfolgsverantwortlichkeit auferlegt.[579] Für solche Techniken werden mehrere Beispiele in Erwägungsgrund 133 S. 4 der KI-VO aufgeführt, wie u. a. Wasserzeichen, Metadatenidentifizierungen oder Protokollierungsmethoden sowie eine Kombination solcher Techniken. Den angeführten Beispielen ist jedoch gemein, dass sie oftmals schwerlich wirksam bzw. nicht auf den vorliegend relevanten Fall der Kennzeichnung des Textes übertragen werden können.[580] Diesen Befund unterstreicht auch die derzeitige Lage in der Praxis. Da Art. 50 KI-VO seit dem 02.08.2026 gilt (für vor diesem Zeitpunkt in Verkehr gebrachte Systeme gilt die Kennzeichnungspflicht nach Abs. 2 aufgrund der Verordnung (EU) 2026/1744 erst ab dem 02.12.2026), mussten sich Anbieter bereits mit dieser Thematik auseinandersetzen. So soll z. B. die Text-Kennzeichnung in den Claude-Modellen des Unternehmens Anthropic mittels eines – bereits auf Modellebene angewendeten – eingebetteten Wasserzeichens erfolgen.[581] Dies funktioniert wie folgt: Bei der Wahl von Wörtern für den Text wählt das System bei mehreren gleichwertigen Möglichkeiten bestimmte Wörter, wodurch sich im Gesamttext ein unsichtbares Muster ergibt.[582] Jedoch ist diese Technik bereits von Haus aus bestimmten Einschränkungen unterworfen, z. B. kann trotz künstlicher Generierung bei einem zu kurzen Text kein zuverlässiges Muster/Signal in diesem implementiert werden.[583] Auch sind bereits kurz nach der Ankündigung mehrere Techniken entwickelt worden, das Wasserzeichen wieder zu entfernen, indem etwa (andere) KI-Modelle in mehreren Durchläufen im Text Wörter durch andere Synonyme austauschen oder indem Texte in andere Sprachen und wieder zurück übersetzt werden.[584] Bei den technischen Lösungen ist (somit) letztlich der allgemein anerkannte Stand der Technik zu berücksichtigen. Bei der Wahl der Mittel ist somit der jeweils gegenwärtige technologische Meinungsstand im Auge zu behalten.[585]
Aufschluss können bei der Wahl der Mittel auch Verzeichnisse über gängige Kennzeichnungsmethoden geben, die daher berücksichtigt und deren Inhalte – soweit vorteilhaft – implementiert werden können.[586] Auch Entwicklungsansätze einheitlicher Standards für maschinenlesbare Kennzeichen bieten Anknüpfungspunkte für die Implementierung in der Praxis.[587] Die konkreten Anforderungen an die Kennzeichnung (z. B. auch die Frage, ob durch sie das Ausmaß der Künstlichkeit bzw. welche Teile künstlich modifiziert wurden) führen zu einer Rechtsunsicherheit und einem

[579] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 289; *Molavi Vasse'i*, RDi 2024, 406, 408.
[580] HK-KI-VO/Hilgendorf/Härtlein/*Hilgendorf/Härtlein*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 11; *Molavi Vasse'i*, RDi 2024, 406, 409f.
[581] *Anthropic*, Wie Claude KI-generierte Inhalte kennzeichnet, abrufbar unter https://support.claude.com/de/articles/16266773-wie-claude-ki-generierte-inhalte-kennzeichnet (Abruf am 01.09.2026, 14:59 Uhr).
[582] *Bölling*, Anthropic erklärt, wie das neue Wasserzeichen in Claude genau funktioniert, abrufbar unter https://t3n.de/news/anthropic-erklaert-wie-das-neue-wasserzeichen-in-claude-genau-funktioniert-1758332/ (Abruf am 01.09.2026, 15:03 Uhr).
[583] *Anthropic*, Wie Claude KI-generierte Inhalte kennzeichnet, abrufbar unter https://support.claude.com/de/articles/16266773-wie-claude-ki-generierte-inhalte-kennzeichnet (Abruf am 01.09.2026, 14:59 Uhr).
[584] *Fuhrmann*, Claude-Wasserzeichen entfernen: Dieses Tool verspricht Abhilfe – aber funktioniert es wirklich?, abrufbar unter https://t3n.de/news/claude-wasserzeichen-entfernen-tool-anthropic-synthid-1758888/ (Abruf am 01.09.2026, 15:11 Uhr).
[585] *Molavi Vasse'i*, RDi 2024, 406, 408 m. w. N.
[586] *Molavi Vasse'i*, RDi 2024, 406, 408 m. H. a. das Community-betriebene Glossar von *Partnerships in* [sic!; korrekt: on] *AI*, Building a Glossary for Synthetic Media Transparency Methods, Part 1: Indirect Disclosure, 2023.
[587] Vgl. *Molavi Vasse'i*, RDi 2024, 406, 409.

anbieterseitigen Konkretisierungsrisiko, das durch Standards nach Art. 50 Abs. 2 S. 2 KI-VO, mit deren Ausarbeitung das European Committee for Standardization sowie das European Committee for Electrotechnical Standardization seitens der Kommission beauftragt wurden, abgemildert werden soll; ebenfalls abmildernd sollen die Leitlinien zur Umsetzung der Transparenzverpflichtungen nach Art. 50 KI-VO nach Art. 96 Abs. 1 lit. d KI-VO wirken, die unlängst, am 20.07.2026, veröffentlicht wurden.[588] Letztlich sind bei der Implementierung von Transparenzinstrumenten allerdings mögliche Kollisionen verschiedener Belange zu beachten: So können etwa in diesem Rahmen Systeme jede Veränderung einer Datei einschließlich Bearbeiter und Zeitstempel tracken, was mit datenschutzrechtlichen Belangen sowie jedenfalls akzeptanztechnisch mit der Richterschaft vor dem Hintergrund der richterlichen Unabhängigkeit nach Art. 97 Abs. 1 GG in Konflikt geraten kann.[589] Bei der Auswahl der Instrumente zur Realisierung der Kennzeichnungspflicht ist insoweit auf eine harmonische Abstimmung zu achten. Die Kennzeichnungspflicht künstlich generierter Inhalte ist ein Regulierungsansatz, der auch in Jurisdiktionen jenseits der europäischen Sphäre gewählt wird, sowie von Betreibern großer Plattformen.[590] Die entsprechenden Regelungen sollten im Blick behalten und die in ihnen niedergelegten Umsetzungstechniken bei Vorteilhaftigkeit für den Einsatz informationstechnischer Systeme im Rahmen der richterlichen Entscheidungsfindung adaptiert werden.

Für den entscheidungs*ersetzenden* Einsatz informationstechnischer Systeme bedeutet das Vorstehende, dass Anbieter hiernach verpflichtet sind, die Herkunft der automatisiert generierten Texte für Menschen und für Maschinen kenntlich zu machen. Dies könnte etwa in der schriftlichen Fassung eines automatisierten Urteils über einen entsprechenden **textlichen Zusatz** geschehen, **dass das Urteil durch ein informationstechnisches System automatisch generiert worden ist**. Auch in einer elektronisch abrufbaren Version könnte in der jeweiligen Datei ein entsprechender, maschinenlesbarer Hinweis oder Code enthalten sein, der den maschinellen Ursprung des Urteils offenlegt.

Eine entsprechende Kennzeichnungspflicht erstreckt sich in gleichem Maße auch auf Systemergebnisse, die im Rahmen des entscheidungs*unterstützenden* Einsatzes generiert worden sind. Auch hier ist ein Hinweis auf die „synthetische“ Erstellung erforderlich. Mit Blick auf dieses Einsatzszenario sind allerdings die **Ausnahmetatbestände** des Art. 50 Abs. 2 S. 3 Var. 1 u. 2 KI-VO zu beachten, deren Anwendung hier diskutabel ist. Hiernach sind bei einer *unterstützenden Funktion für die Standardbearbeitung* und der *nicht wesentlichen Veränderung* der vom Betreiber *bereitgestellten Eingabedaten oder deren Semantik*, für den Einsatz informationstechnischer Systeme keine Kennzeichnungsnotwendigkeiten gegeben.

---

[588] BeckOK KI-Recht/*Lauber-Rönsberg*, 7. Ed. 25.7.2026, KI-VO Art. 50 Rn. 35f.; *Molavi Vasse'i*, RDi 2024, 406, 413; *Europäische Kommission,* Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme nach Artikel 50 der KI-Verordnung „Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the 'AI Act')“, abrufbar unter https://digital-strategy.ec.europa.eu/de/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems (Abruf am 31.08.2026, 20:51 Uhr) sowie die Pressemitteilung zur Veröffentlichung, *Europäische Kommission*, Pressemitteilung 20.07.2026, Kommission veröffentlicht Leitlinien zu Transparenzpflichten für Anbieter und Betreiber bestimmter KI-Systeme, abrufbar unter https://ec.europa.eu/commission/presscorner/detail/de/ip_26_1653 (Abruf am 31.08.2026, 20:51 Uhr).

[589] Dies zu den Anforderungen des Datenschutzes anführend *Molavi Vasse'i*, RDi 2024, 406, 414.

[590] *Molavi Vasse'i*, RDi 2024, 406, 406 mit Hinweisen auf einschlägige Regelungen.

Diese Ausnahmebestimmungen sind generell restriktiv[591] zu handhaben und kommen angesichts der Tatbestandsgestaltung nur bei marginalen bzw. kosmetischen Änderungen in Betracht.[592] Der ratio legis nach können unter Art. 50 Abs. 2 S. 3 KI-VO damit nur Systeme mit Funktionalitäten fallen, die den Informationsgehalt eines Textes nicht grundlegend verändern.[593]
Die Ausnahmen kommen etwa bei rein administrativen Funktionalitäten wie der Metadatenextraktion in Betracht (unterstützende Funktion für die Standardbearbeitung). Auch kommt die Anwendung des Ausnahmetatbestands dann in Betracht, wenn ein KI-System ferner lediglich Vorschläge zur Verbesserung des Schreibstils unterbreitet (keine wesentliche Veränderung der Semantik).[594] Bereits Funktionalitäten wie die Übersetzung von fremdsprachigem juristischem Text werden allerdings außerhalb des Anwendungsbereichs der Ausnahmetatbestände liegen, da hier u. U. Rechtsphänomene übersetzt werden müssen, deren Inhalt und Bedeutung von Rechtsordnung zu Rechtsordnung differieren und somit einer präzisen Wiedergabe und Erfassung bedürfen.[595] Semantische Feinheiten sind hier von erheblicher Bedeutung und aufgrund der gegenwärtigen limitierten Leistungsfähigkeit von KI-Systemen kann semantische Präzision nicht sichergestellt werden. Eine Kennzeichnungspflicht ist daher notwendig. Zwar weisen die veröffentlichten (unverbindlichen) Leitlinien (zu Leitlinien siehe noch nachfolgend im Abschnitt C.II.4.b.aa(2)(iv)) zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme nach Art. 50 der KI-Verordnung der Europäischen Kommission die Übersetzung von Text als mögliches Beispiel für die Einschlägigkeit der Ausnahmen nach Art. 50 Abs. 2 S. 3 KI-VO aus, anders als noch deren Entwurfsfassung, bei der die Übersetzung von Text diametral entgegengesetzt noch als Beispiel für semantische Veränderungen aufgeführt wurde, die eine Kennzeichnung nach Art. 50 Abs. 2 KI-VO erfordern.[596] Das Beispiel wird jedoch nur pauschal gegeben ohne Differenzierung nach Einsatzgebiet (z. B. Übersetzung von Alltagstexten wie Songtexten vs. Übersetzung von ausländischen juristischen Texten). Zudem lässt sich dem systematischen Vergleich mit den weiteren gegebenen Beispielen entnehmen, dass diese, entsprechend der Ausnahmeregelung, beschränkt auf minimalinvasive Maßnahmen zu verstehen sind, die den Fall der Übersetzung fremdsprachiger juristischer Texte aus den vorstehend genannten Gründen nicht umfassen. Besonders deutlich lässt sich dies anhand folgenden gegebenen Beispiels für die Einschlägigkeit

---

[591] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 289; siehe auch Bomhard/Pieper/Wende/*Merkle*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 107, 110ff.
[592] So bereits m. w. N. *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 289f.
[593] Bomhard/Pieper/Wende/*Merkle*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 110.
[594] BeckOK KI-Recht/*Lauber-Rönsberg*, 7. Ed. 25.7.2026, KI-VO Art. 50 Rn. 38.
[595] Vgl. so auch noch der Entwurf der Kommission für Leitlinien zur Umsetzung der Transparenzpflichten nach Art. 50 KI-VO, *Europäische Kommission*, Entwurf der Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme gemäß Artikel 50 der KI-VO, abrufbar unter https://digital-strategy.ec.europa.eu/en/library/draft-guidelines-implementation-transparency-obligations-certain-ai-systems-under-article-50-ai-act (Abruf am 13.05.2026, 16:15 Uhr), S. 22f.
[596] Veröffentlichte Leitlinien: *Europäische Kommission,* Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme nach Artikel 50 der KI-Verordnung „Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the ‘AI Act’)“, abrufbar unter https://digital-strategy.ec.europa.eu/de/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems (Abruf am 31.08.2026, 20:51 Uhr), Rn. 92; Entwurfsfassung Leitlinien: *Europäische Kommission*, Entwurf der Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme gemäß Artikel 50 der KI-VO, abrufbar unter https://digital-strategy.ec.europa.eu/en/library/draft-guidelines-implementation-transparency-obligations-certain-ai-systems-under-article-50-ai-act (Abruf am 13.05.2026, 16:15 Uhr), S. 22f.

der Ausnahmeregelung sehen: KI-generierte Inhalte, die authentische menschliche Eingaben durch unterstützende Technologien lediglich umwandeln, zur Ermöglichung von Kommunikation für behinderte Personen. Die Einschlägigkeit der Ausnahmeregelung wird in diesem Fall damit begründet, dass diese Technologien nicht die Bedeutung des Inhalts semantisch verändern.[597] Auch in diesen Fällen geht es, wie bei der Übersetzung von fremdsprachigem juristischem Text, um die Übersetzung von Eingaben. Mit Blick auf die immanente Gefahr unpräziser Wiedergabe und Erfassung sowie die fehlende Sicherstellung semantischer Präzision, wäre die Kategorisierung der Übersetzung ausländischer juristischer Texte auch mit dem Telos der Ausnahmeregelungen des Art. 50 Abs. 2 S. 3 KI-VO unvereinbar. Sie ist (damit) als nicht den Ausnahmeregelungen unterfallend anzusehen.
Für weitergehende sachlich-inhaltliche Funktionalitäten wie das Generieren von Text, wie es beispielsweise das Erstellen eines Tatbestandsentwurfs, die Abfassung einer Relationstabelle mit sinngemäßer Zusammenfassung und Wiedergabe von Akteninhalt oder gar die Vorbereitung von ganzen Urteilsentwürfen erfordern, sind die Ausnahmetatbestände nicht einschlägig. Es ist daher regelmäßig eine Kennzeichnungsnotwendigkeit gegeben. Einzig dann, wenn das KI-System **vom menschlichen Nutzer selbst markierten Text** textlich unverändert übernimmt und in einen Tatbestandsentwurf überträgt (z. B. vom Richter mit Markierfunktion gekennzeichneten Text in einen Tatbestandsentwurf an die Stelle eines vorgesehenen Platzhalters unverändert überträgt), wäre an eine Einschlägigkeit des Ausnahmetatbestands zu denken, da die semantische Präzision dann sichergestellt wäre.
In der Praxis werden sich bzgl. der Ausnahmetatbestände erhebliche Abgrenzungsschwierigkeiten ergeben. Es ist davon auszugehen, dass Anbieter zur Vermeidung von Schäden infolge von Sanktionen bei Nicht-Einschlägigkeit der Ausnahmeregelungen sowie bezüglich ihrer Reputation **im Zweifel die Kennzeichnung** vornehmen werden.[598] Dies ist auch im Hinblick auf den entscheidungs*unterstützenden* Einsatz **durchaus sinnvoll**, da Richterinnen und Richtern insoweit bei der Nutzung von Unterstützungssystemen stets verdeutlicht wird, dass entsprechender Text von einem KI-System generiert worden ist. Dies kann auch bei einer über mehrere Tage, Wochen oder sogar Monate hinweg erfolgenden Arbeit mit dem Akteninhalt hilfreich sein, automatisiert generierten Text auch zu einem späteren Zeitpunkt noch als solchen zu erkennen.
Mit Blick auf das **„Wie“ der Kennzeichnung** bietet sich neben den textlichen Hinweisen, dass ein Text automatisiert von einem KI-System generiert worden ist, an, den generierten Text als solchen auch optisch von sonstigem Text abzuheben, so dass die automatisiert generierten Teile von dem Richter bzw. der Richterin leicht erkannt und nachvollzogen werden können (z. B. über **farbliche Hervorhebung bzw. Umrahmung**).
Fraglich ist, wie weit die Kennzeichnungspflicht von Anbietern insoweit reicht (**Reichweite der Kennzeichnungspflicht**). Konkret ist fraglich, ob „synthetisch“ generierter Text beispielsweise auch dann noch als solcher gekennzeichnet werden muss, wenn er durch Nutzerinnen oder Nutzer weiterverarbeitet worden ist. Das wäre beispielsweise der Fall, wenn Richterinnen und Richter sich durch ein

[597] *Europäische Kommission,* Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme nach Artikel 50 der KI-Verordnung „Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the ‘AI Act’)“, abrufbar unter https://digital-strategy.ec.europa.eu/de/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems (Abruf am 31.08.2026, 20:51 Uhr), Rn. 92.
[598] Bomhard/Pieper/Wende/*Merkle*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 109.

Unterstützungssystem Textteile für ein Urteil generieren lassen (z. B. den Tatbestand abfassen lassen) und diesen Tatbestand dann im finalen Urteil konkret angepasst und weiterverarbeitet haben. Die Frage lautet also konkret: Muss das fertige Urteil dann die KI-generierten Textteile als solche kenntlich ausweisen und muss dies der Anbieter sicherstellen (Anbieterpflicht)? Hierfür spricht, dass die Anbieterpflichten nach Art. 50 Abs. 2 KI-VO und die Betreiberpflichten nach Art. 50 Abs. 4 KI-VO zunächst grundsätzlich nebeneinander stehen und der Regelungsgedanke der Anknüpfung des Abs. 2 an den Anbieter ist, dass der Ansatz bei ihnen, am technischen Ursprung, am vielversprechendsten für eine einheitliche sowie verlässliche Kennzeichnung ist.[599] Zudem beinhaltet Art. 50 Abs. 2 S. 3 KI-VO ausdrücklich geregelte Ausnahmen von der Kennzeichnungspflicht, die für den vorliegenden Fall nicht einschlägig sind.[600] Gleichwohl ist die Frage aus mehreren Gründen verneinend zu beantworten: Erstens fällt die Weiterverarbeitung des KI-generierten Texts nicht mehr in die Verantwortungssphäre des Anbieters, da er keinerlei Einfluss darauf hat, was die Nutzenden mit dem generierten Text final tun. Er hat daher auch keinen Einfluss darauf, dass die jeweiligen Kennzeichnungen im vom menschlichen Nutzer final angepassten Ergebnis enthalten sind. Er kann dies weder sicherstellen noch darauf Einfluss nehmen, dass die Pflicht erfüllt wird. Auch wird in der Literatur ausgeführt, dass die KI-VO der betreiberseitigen Löschung der Anbieter-Kennzeichnung in Fällen, in denen keine Offenlegungspflicht nach Art. 50 Abs. 4 KI-VO besteht, nicht entgegensteht.[601] Es ist daher nicht sinnvoll, dem Anbieter diese Pflicht aufzuerlegen, da er ihre Einhaltung nicht sicherstellen kann. Jemandem vor dem Hintergrund nicht kontrollierbarer Pflichteinhaltung ggf. sanktionsbewehrt eine Pflicht aufzuerlegen wäre darüber hinaus rechtspolitisch höchst zweifelhaft. Diese Problematik erkennt letztlich auch die KI-VO, indem sie in Art. 50 Abs. 4 KI-VO dem Betreiber eines KI-Systems eine eigene Pflicht zur Offenlegung des Einsatzes von KI-Systemen bei der Textgenerierung auferlegt. Darüber hinaus ermöglicht Art. 50 Abs. 4 UAbs. 2 S. 2 KI-VO mit Blick auf die Betreiberpflicht zur Offenlegung KI-generierten Inhalts ebenfalls eine Ausnahme, wenn der KI-generierte Inhalt von einer natürlichen Person im Nachgang überprüft worden ist und diese Person (oder eine andere) die (redaktionelle) Verantwortung für diesen Text übernommen hat (dazu sogleich). Die Verantwortungszuweisung liegt in diesen Fällen bei der die Überprüfung vornehmenden Person, d. h. im richterlichen Entscheidungskontext, bei dem/der erkennenden Richter/in, der/die das Urteil erlassen hat. Der Betreiber hat den KI-Einsatz in diesen Fällen nicht offenzulegen. Die Anbieterpflicht kann insoweit aber nicht weitergehen als die Betreiberpflicht, sodass die Kennzeichnungspflicht sich letztlich bei Unterstützungssystemen bei der richterlichen Entscheidungsfindung nicht auf das vom menschlichen Richter bzw. der menschlichen Richterin final erlassene Urteil erstreckt. Dies würde zudem dazu führen, dass der Anbieter sanktionsbewehrt verpflichtet wäre, eine grundsätzlich rechtmäßige Maßnahme des Betreibers, in Form der Entfernung der Kennzeichnung bei fehlender Offenlegungspflicht, zu verhindern. Mit Blick auf die Ausnahmeregelungen in Art. 50 Abs. 2 S. 3 KI-VO sowie Art. 50 Abs. 4 UAbs. 2 S. 2 KI-VO kann zudem angeführt werden, dass beide dem Vorgehen gegen synthetische Desinformation dienen, was insoweit auch für den Gleichlauf von Anbieter- und Betreiberpflicht angebracht werden kann.[602] Eine

[599] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 50 Rn. 70.
[600] Vgl. zu den Ausnahmen etwa Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 50 Rn. 82ff.; BeckOK KI-Recht/*Lauber-Rönsberg*, 7. Ed. 25.7.2026, KI-VO Art. 50 Rn. 32.
[601] BeckOK KI-Recht/*Lauber-Rönsberg*, 7. Ed. 25.7.2026, KI-VO Art. 50 Rn. 32 i. V. m. 67.
[602] Vgl. *Oster*, ZfDR 2026, 163, 172.

Fortsetzung der Kennzeichnungspflicht im weiterverarbeiteten Ergebnis ist damit abzulehnen und lässt sich der Norm nicht entnehmen.

(iii) Offenlegungspflicht „synthetisch“ generierter Textinhalte (Betreiberpflicht)

Art. 50 Abs. 4 UAbs. 2 S. 1 KI-VO beinhaltet ergänzend eine grundsätzliche **Offenlegungspflicht KI-generierter Inhalte für Betreiber von KI-Systemen, die Texte erzeugen** oder manipulieren, wenn der erzeugte oder manipulierte Text veröffentlicht wird, **um die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse zu informieren**. Im nationalen Rechtskreis betreffend die richterliche Entscheidungsfindung korrespondiert hiermit das Informationsinteresse der Bürgerinnen und Bürger, zur Kontrolle des Staates in Form der Rechtsprechung.[603] Zweck der Offenlegungspflicht des Art. 50 Abs. 4 UAbs. 2 S. 1 KI-VO ist der **Schutz der Authentizität des öffentlichen Meinungskampfes**, des **Vertrauens in die Integrität von meinungsbildungsrelevanten Inhalten**.[604] Entsprechend ist auch nicht jede Veröffentlichung von künstlich generiertem Text erfasst, sondern meinungsbildungsrelevante Inhalte („*um die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse zu informieren*“), worunter u. a. politische und gesellschaftliche Angelegenheiten, auch nur mit lokaler oder partikularer Relevanz fallen.[605] Der Tatbestand der „Öffentlichkeit“ ist erfüllt, wenn der Diskursraum grundsätzlich jedermann eröffnet ist, nicht hingegen, wenn nach der hierfür relevanten Veröffentlichungsintention nur ein überschaubarer oder abgegrenzter Personenkreis erreicht werden kann und soll.[606]
Die Vorgaben des Art. 50 Abs. 4 UAbs. 2 S. 1 KI-VO werden insbesondere im Zusammenhang mit der Veröffentlichung von gerichtlichen Entscheidungen relevant, die unter Verwendung informationstechnischer Systeme gefällt wurden. Dies gilt für Entscheidungen, die durch entscheidungs*ersetzende* informationstechnische Systeme gefällt werden, ebenso wie für solche im entscheidungs*unterstützenden* Einsatzszenario. Gerichtliche Entscheidungen, insbesondere obergerichtlichen oder höchstrichterlichen Ursprungs, befassen sich oftmals mit gesellschaftlich relevanten Themen und werden medial eng begleitet. Aber auch sonstige gerichtliche Entscheidungen, deren flächendeckende Veröffentlichung sukzessive forciert wird, beschäftigen sich mit (zwischen den Parteien) streitbefangenen Themen, deren Relevanz oftmals über die Parteien des Streits hinausstrahlt und zudem auch mitunter medial begleitet werden. Veröffentlichte Entscheidungen unterfallen der Kategorie der meinungsbildenden Inhalte.[607] Dies ergibt sich beispielsweise aus der grundsätzlichen Funktion der Öffentlichkeit gerichtlicher Verfahren, die eine Überprüfung der Justiz durch die Öffentlichkeit ermöglichen, vor dem Hintergrund nationalsozialistischen Unrechts eine Geheimjustiz nachgerade verhindern soll und letztlich dem

[603] Siehe hierzu nur *BVerwG*, NJW 1997, 2694, 2695; *BVerfGE* 103, 44, 63; *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 137f.; Dürig/Herzog/Scholz/*Mehde*, 111. EL Juni 2026, GG Art. 28 Rn. 57; Dürig/Herzog/Scholz/*Grzeszick*, 111. EL Juni 2026, GG Art. 20 Rn. 34; *Reuß*, PdR 2026, 77, 86f.
[604] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 50 Rn. 143.
[605] BeckOK KI-Recht/*Lauber-Rönsberg*, 7. Ed. 25.7.2026, KI-VO Art. 50 Rn. 65; Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 50 Rn. 149.
[606] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 50 Rn. 148.
[607] Vgl. so auch *Europäische Kommission,* Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme nach Artikel 50 der KI-Verordnung „Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the ‘AI Act’)“, abrufbar unter https://digital-strategy.ec.europa.eu/de/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems (Abruf am 31.08.2026, 20:51 Uhr), Rn. 131.

demokratischen Informationsinteresse der Bevölkerung, sich über die Vorgänge der Ausübung staatlicher Gewalt zu informieren, gerecht wird. Der Inhalt von Gerichtsentscheidungen ist somit stets von öffentlichem Interesse. Die Öffentlichkeit ist durch die Veröffentlichung der Entscheidungen stets tangiert. Dass die Entscheidungen indes ggf. hinter einer Bezahlwand oder erst infolge eines abgeschlossenen Abonnements zugänglich sein könnten, lässt dies unberührt.[608]
Art. 50 Abs. 4 UAbs. 2 S. 1 KI-VO verlangt eine Offenlegung der künstlichen Erzeugung bzw. Manipulation der Inhalte; Ziel ist, dass die KI-Generierung derart nach außen ausgewiesen wird, dass dies von allen Rezipienten wahrgenommen wird.[609] Anders als in Art. 50 Abs. 2 KI-VO verlangt Art. 50 Abs. 4 KI-VO keine Maschinenlesbarkeit.[610] Wie die Offenlegung zu erfolgen hat, spezifiziert Art. 50 Abs. 4 KI-VO nicht, vorgeschlagen wird jedoch eine Anlehnung an Art. 50 Abs. 2 KI-VO, dass die künstliche Erzeugung oder Manipulation für den Menschen erkennbar ist.[611] ErwGr. 134 S. 4 i. V. m. 1 benennt für die Offenlegung eine Kennzeichnung der KI-Ausgaben und Hinweise auf ihren künstlichen Ursprung bzw. dem ähnliche Maßnahmen. Die Offenlegung muss hierbei den Vorgaben des Art. 50 Abs. 5 KI-VO entsprechen, was bedeutet, dass sie spätestens zum Zeitpunkt der ersten Interaktion oder Aussetzung in klarer und eindeutiger Weise zu erfolgen hat, wobei die Informationen den geltenden Barrierefreiheitsanforderungen entsprechen müssen. Hierbei sollte auf eine hohe Wahrnehmbarkeit, leichte Verständlich- und Zugänglichkeit sowie Abgrenzbarkeit der Information von anderen sowie der sonstigen Darstellungsumgebung geachtet werden.[612] Als Maßnahmen hierfür werden u. a. vorgeschlagen gut erkenntliche Sticker „KI-generiert“ oder Icons sowie Über- oder Unterschriften bzw. einleitende oder abschließende Hinweise und Wasserzeichen, die auf den KI-Ursprung hinweisen.[613] In den veröffentlichten Entscheidungen können zur Umsetzung der Offenlegungspflicht **standardmäßige Hinweise an festgelegten Stellen**, wie etwa nach dem Tenor und vor dem Tatbestand oder am Ende der Entscheidung platziert und bei Einsatz informationstechnischer Systeme automatisiert eingefügt werden. Auch ist insbesondere beim Einsatz entscheidungsersetzender informationstechnischer Systeme bei der richterlichen Entscheidungsfindung die Aufnahme des Hinweises im Rubrum denkbar und könnte einheitlich mit den o. g. Anbieterpflichten systemseitig bereits vorgesehen werden.

---

608 Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 50 Rn. 148; vgl. so auch *Europäische Kommission,* Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme nach Artikel 50 der KI-Verordnung „Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the ‘AI Act’)“, abrufbar unter https://digital-strategy.ec.europa.eu/de/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems (Abruf am 31.08.2026, 20:51 Uhr), Rn. 131.

609 BeckOK KI-Recht/*Lauber-Rönsberg*, 7. Ed. 25.7.2026, KI-VO Art. 50 Rn. 69; HK-KI-VO/Hilgendorf/Härtlein/*Hilgendorf/Härtlein*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 17, 13ff.

610 *Oster*, ZfDR 2026, 163, 180 i. V. m. 176.

611 *Oster*, ZfDR 2026, 163, 180 i. V. m. 177; vgl. zu einem Gleichlauf der Anforderungen auch Bomhard/Pieper/Wende/*Merkle*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 73.

612 *Europäische Kommission,* Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme nach Artikel 50 der KI-Verordnung „Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the ‘AI Act’)“, abrufbar unter https://digital-strategy.ec.europa.eu/de/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems (Abruf am 31.08.2026, 20:51 Uhr), Rn. 142.

613 HK-KI-VO/Hilgendorf/Härtlein/*Hilgendorf/Härtlein*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 18; Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 50 Rn. 150; Bomhard/Pieper/Wende/*Merkle*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 161.

Im Rahmen einzelner künstlich generierter Textteile beim entscheidungs*unterstützenden* Einsatz bieten sich Icon-Kennzeichnungen[614] an den entsprechenden Stellen in der Entscheidung an, mit Blick auf eine möglichst exakte Nachvollziehbarkeit des Einsatzes informationstechnischer Systeme. Im Rahmen von gerichtlichen Hinweisen oder Mitteilungen im Rahmen des gerichtlichen Verfahrens dürfte die Offenlegungspflicht hingegen mangels Vorliegens von „Öffentlichkeit“ i. S. d. Vorschrift nicht greifen. Derartige Schriftsätze ergehen im Verfahren an die Parteien und ggf. deren Vertreter, mithin einen klar abgegrenzten Personenkreis. Eine weitere Verbreitung über diesen Personenkreis ist nicht intendiert. Auch werden diese Schriftsätze nach dem regelmäßigen Gang der Dinge nicht öffentlich zugänglich. Sollte jedoch im Rahmen der Digitalisierung des Zivilprozesses die vollständige Veröffentlichung der Prozessakten forciert werden, im Sinne einer Stärkung des Öffentlichkeitsgrundsatzes,[615] würden auch diese den Tatbestand der „Öffentlichkeit“ erfüllen und müssten entsprechend gekennzeichnet werden, sofern nicht die Ausnahmetatbestände greifen.
Art. 50 Abs. 4 UAbs. 2 S. 2 KI-VO enthält von dieser Offenlegungspflicht Ausnahmen. Im vorliegenden Betrachtungsspektrum relevant sind der Fall, dass die durch KI erzeugten Inhalte einem **Verfahren der menschlichen Überprüfung** oder **redaktionellen Kontrolle** unterzogen wurden und dass **eine natürliche oder juristische Person die redaktionelle Verantwortung für die Veröffentlichung** der Inhalte **trägt**. Zur Verwirklichung des Ausnahmetatbestandes bedarf es **kumulativ** der Erfüllung beider Voraussetzungen: erstens muss ein Verfahren der menschlichen Überprüfung oder redaktionellen Kontrolle vorliegen und zweitens muss eine natürliche oder juristische Person die redaktionelle Verantwortung für die Veröffentlichung der Inhalte tragen.[616] Den Varianten der ersten Voraussetzung ist gemein, dass sie eine inhaltliche, menschliche Überprüfung bezüglich der Richtigkeit der Inhalte des KI-manipulierten bzw. -generierten Textes erfordern.[617] Die Intensität der Kontrolle spezifiziert die KI-VO nicht.[618] Die Pflicht zur Aufsicht und Kontrolle des Systems und seiner Ausgaben ergibt sich bereits aus einer Vielzahl weiterer rechtlicher Verbürgungen (s. o., insbesondere Abschnitt C.II.2.a).

---

[614] Bomhard/Pieper/Wende/*Merkle*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 161.
[615] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 351 m. w. N.
[616] So die herrschende Ansicht in der Literatur: BeckOK KI-Recht/*Lauber-Rönsberg*, 7. Ed. 25.7.2026, KI-VO Art. 50 Rn. 73; Bomhard/Pieper/Wende/*Merkle*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 163; HK-KI-VO/Hilgendorf/Härtlein/*Hilgendorf/Härtlein*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 29; so auch *Europäische Kommission,* Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme nach Artikel 50 der KI-Verordnung „Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the ‘AI Act’)“, abrufbar unter https://digital-strategy.ec.europa.eu/de/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems (Abruf am 31.08.2026, 20:51 Uhr), Rn. 133; für die Ausnahme die Erfüllung nur eines der Elemente wohl als ausreichend erachtend Martini/Wendehorst/*Martini*, 1. Aufl. 2024, KI-VO Art. 50 Rn. 114, worauf auch BeckOK KI-Recht/*Lauber-Rönsberg*, 7. Ed. 25.7.2026, KI-VO Art. 50 Rn. 73 hinweist. Diese Ansicht scheint in der Neuauflage von Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 50 Rn. 152 nicht mehr vertreten zu werden.
[617] BeckOK KI-Recht/*Lauber-Rönsberg*, 7. Ed. 25.7.2026, KI-VO Art. 50 Rn. 73; Bomhard/Pieper/Wende/*Merkle*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 164; HK-KI-VO/Hilgendorf/Härtlein/*Hilgendorf/Härtlein*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 29; so auch *Europäische Kommission,* Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme nach Artikel 50 der KI-Verordnung „Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the ‘AI Act’)“, abrufbar unter https://digital-strategy.ec.europa.eu/de/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems (Abruf am 31.08.2026, 20:51 Uhr), Rn. 134f.
[618] Bomhard/Pieper/Wende/*Merkle*, 1. Aufl. 2025, KI-VO Art. 50 Rn. 164.

Die **inhaltliche Überprüfung der Ergebnisse** obliegt im Rahmen der richterlichen Entscheidungsfindung im entscheidungs*unterstützenden* Einsatz ohnehin dem **nutzenden Richter bzw. der nutzenden Richterin**, der/die **zugleich die Letztverantwortung** trägt. Es erfolgt damit ohnehin eine Kontrolle der Ergebnisse des Systems. Deren Intensität hängt von dem funktionalen Charakter der Tätigkeit, ihrem Automatisierungsgrad sowie der Komplexität des jeweiligen Anwendungsfeldes ab. Insoweit wird der ersten Voraussetzung der Ausnahme im Rahmen des entscheidungs*unterstützenden* Einsatzes Rechnung getragen.
Im Rahmen des entscheidungs*ersetzenden* Einsatzes informationstechnischer Systeme treffen diese die Entscheidung selbst. Inwieweit eine auch für die Ausnahme zur Offenlegungspflicht inhaltliche Kontrolle der Ergebnisse gegeben ist, hängt maßgeblich von der konkreten Ausgestaltung des entscheidungsersetzenden Einsatzes ab. Soweit **keine (vollständige) individuelle inhaltliche Überprüfung der jeweiligen Entscheidung** erfolgt, dürfte die Voraussetzung des Ausnahmetatbestandes nicht erfüllt sein und es müsste damit die Offenlegung der künstlichen Texterzeugung oder -manipulation erfolgen. Soweit etwa eine **stichprobenartige Überprüfung** einer Entscheidung des Inhalts erfolgt und ihr Inhalt durch einen menschlichen Richter überprüft wird, könnte hingegen die erste Voraussetzung dieses Ausnahmetatbestands zur Offenlegungspflicht für das stichprobenartig überprüfte Verfahren erfüllt sein. Eine generalisierte Offenlegung wird aber bereits deshalb angezeigt sein, da die **Überprüfung meist ex post** erfolgen wird, so dass **im Zeitpunkt des Ergehens der Entscheidung der Ausnahmetatbestand noch nicht erfüllt** ist. Somit ist im **Regelfall eine Offenlegung** durch den Betreiber, der mit der Kennzeichnung des Anbieters gleichzeitig systemseitig erfüllt werden kann und sollte, als **erforderlich** angesehen werden.
Als zweite Voraussetzung muss eine natürliche oder juristische Person die **redaktionelle Verantwortung für die Veröffentlichung** der Inhalte tragen. Dies beschreibt, dass Verantwortung für den künstlich generierten Text übernommen wird.[619] Unter Bezugnahme auf medienrechtliche Begrifflichkeiten wird angeführt, dies würde dadurch hergestellt, dass eine natürliche Person den Textinhalt sichtet, auswählt und mit Bearbeitungsbefugnis prüft. Zudem wird eine Festlegung im Rahmen einer redaktionellen Entscheidung gefordert hinsichtlich der Modalitäten der Veröffentlichung, Zusammenstellung sowie Bereitstellung an die Allgemeinheit.[620] Angesichts der Wertung des Unionsrechtsgebers, dass das zu eigen machen nach der Kontrolle des künstlich erzeugten oder manipulierten Textes durch einen Menschen die Gefahr für die öffentliche Meinungsbildung durch künstliche, ungeprüfte Information in erheblichem Umfang, hinreichend beschränkt, sollte nicht starr an den medienrechtlichen Begrifflichkeiten festgehalten werden. Von entscheidender Bedeutung ist die Verantwortungsübernahme für den veröffentlichten Text bzw. dessen Inhalt anhand der Wertung des Unionsrechtsgebers.[621] Freilich ist die Ausnahme augenfällig primär mit Blick auf die journalistische Tätigkeit verfasst worden, diese wird präsumtiv auch den Primäranwendungsfall bilden. Jedoch sollte mit Blick darauf, dass nicht nur klassische journalistische Werke, sondern auch andere

[619] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 50 Rn. 152.
[620] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 50 Rn. 152.
[621] Ebenfalls nur auf die Verantwortungsübernahme für den Text abstellend *Europäische Kommission,* Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme nach Artikel 50 der KI-Verordnung „Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the ‘AI Act’)“, abrufbar unter https://digital-strategy.ec.europa.eu/de/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems (Abruf am 31.08.2026, 20:51 Uhr), Rn. 138.

Inhalte, wie etwa Gerichtsentscheidungen von Art. 50 Abs. 4 UAbs. 2 S. 1 KI-VO erfasst sind, die Ausnahmeregelung vor dem Hintergrund ihres Telos nicht ungebührlich im Anwendungsbereich enggeführt werden. Jedenfalls mit Blick auf Gerichtsentscheidungen sollte beachtet werden, dass diese nach Erlass (dem Grunde nach, vgl. §§ 319 ff. ZPO) nicht noch einmal überarbeitet bzw. korrigiert werden und zudem anders als journalistische Werke nicht erst mit ihrer Veröffentlichung an die Allgemeinheit das erste Mal Wirkung nach außen zeitigen. Dies erfolgt vielmehr bereits mit Erlass der Entscheidung gegenüber den Parteien. Die Veröffentlichung des Urteils gegenüber der Allgemeinheit kann insoweit nicht derart erfolgen, dass am Entscheidungsinhalt Änderungen im Sinne einer medienrechtlichen Redaktionstätigkeit vorgenommen werden. Im Wesentlichen ist die geforderte Redaktionstätigkeit vielmehr dem Entscheidungserlass (bei der Tätigkeit menschlicher Richter und Richterinnen) vorgelagert und damit bereits im Stadium der Entscheidungsverkündung abgeschlossen. Dass den Ausnahmeregelungen das Erfassen hoheitlich verfasster Werke nicht fremd ist, zeigt auch Art. 50 Abs. 4 UAbs. 2 S. 2 Alt. 1 KI-VO. Eine Engführung der Ausnahme des Art. 50 Abs. 4 UAbs. 2 S. 2 Alt. 2 KI-VO anhand medienrechtlicher Spezialitäten würde dem Telos nicht gerecht, jedenfalls mit Blick auf gerichtliche Entscheidungen und deren Besonderheiten (bei der Veröffentlichung). Es ist – nach Maßgabe des Art. 50 Abs. 4 UAbs. 2 KI-VO – letztlich auch nicht ersichtlich, weshalb Journalisten mit KI-Unterstützung verfasste Texte nicht aufdecken sollen müssen, dies jedoch stets im Rahmen der Veröffentlichung von Gerichtsentscheidungen erfolgen solle. Es ist nämlich gerade nicht Zweck der Pflicht zur Offenlegung, zu enttarnen, dass ein Text mithilfe von KI erstellt wurde.[622] Mithin **genügt die Verantwortungsübernahme für den (Entscheidungs)Text(inhalt) nach außen für den Tatbestand der redaktionellen Verantwortung** (im Fall der Veröffentlichung gerichtlicher Entscheidungen).

Mit Blick auf den entscheidungs*unterstützenden* Einsatz informationstechnischer Systeme bei der richterlichen Entscheidungsfindung lässt sich damit festhalten, dass aufgrund der den **Richtern ohnehin obliegenden Pflicht zur Prüfung des Inhalts sowie Letztverantwortung bezüglich der Entscheidung**, der **Tatbestand der „redaktionellen Verantwortung“ als erfüllt** angesehen werden kann. Entscheidungen werden mit Hinweis auf das Gericht und ggf. die Kammer bzw. den Senat veröffentlicht, sowie bisweilen sogar unter Nennung der an der Entscheidung beteiligten Richterinnen und Richter, wie bei Entscheidungen des Bundesverfassungsgerichts üblich. Dies dürfte der Voraussetzung genügen, dass eine natürliche Person (der Richter bzw. die Richterin) oder eine juristische Person (in Betracht kämen das Land oder der Bund) die „redaktionelle Verantwortung“ für die Veröffentlichung der Inhalte gem. Art. 50 Abs. 4 UAbs. 2 S. 2 Alt. 2 KI-VO trägt.[623]

Auf den Anwendungsfall des entscheidungs*ersetzenden* Einsatzes übertragen, fehlt nach den zuvor ausgeführten Maßstäben (in der Regel) die Anbindung an die menschliche Kontrolle (vgl. Telos der Ausnahme) bzw. deren Anbindbarkeit im

---

[622] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 50 Rn. 152.

[623] Allgemein für Veröffentlichungen der Justiz von der Einschlägigkeit des Art. 50 Abs. 4 UAbs. 2 S. 2 KI-VO ausgehend Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 24.

Nachgang des Erlasses der Entscheidung für die Verwirklichung des Tatbestandes der „redaktionellen Verantwortung“. Damit ist hier **in der Regel von einer Offenlegungspflicht** auszugehen.
Damit lässt sich festhalten, dass im **entscheidungs*ersetzenden* Einsatzszenario regelmäßig Offenlegungspflichten** des Betreibers greifen und eine Ausnahmeregelung in der Regel nicht durchgreifen wird. Sinnvollerweise sind diese Offenlegungspflichten gemeinsam mit der Kennzeichnungspflicht des Anbieters (dazu bereits oben) durch einen entsprechenden textlichen und maschinenlesbaren Hinweis, dass die Entscheidung automatisiert durch ein KI-System generiert worden ist, zu erfüllen.
Im **entscheidungs*unterstützenden* Einsatzszenario** greift grundsätzlich bei Gerichtsentscheidungen ebenfalls eine **Offenlegungspflicht,** die allerdings **meist durch die richterliche Endkontrolle und Verantwortungsübernahme mittels Erlass der Entscheidung unter den Ausnahmetatbestand zu fassen** ist. Hier ist eine Offenlegung mit Blick auf die Verbürgungen der KI-VO nicht erforderlich. Gleichwohl ist mit Blick auf die Transparenzerfordernisse des Grundgesetzes, die von der KI-VO nicht berührt werden, vgl. Art. 50 Abs. 6 KI-VO, allerdings auch hier eine Information über das „Ob“ und das „Wie“ des KI-Einsatzes zu empfehlen. Dazu eingehend bereits oben.

(iv) Konkretisierungsinstrumente für die Pflichten nach Art. 50 ff. KI-VO

Die KI-VO sieht eine Mehrzahl an Konkretisierungsinstrumenten zur praxisgerechten Umsetzung ihrer Vorgaben vor, namentlich harmonisierte (technische) Normen (Art. 40 ff. KI-VO), Praxisleitfäden (Art. 56 KI-VO), Verhaltenskodizes (Art. 95 KI-VO) sowie Leitlinien (Art. 96 KI-VO).[624] (Derzeit) Relevant sind insbesondere die Leitlinien[625], die Praxisleitfäden[626] sowie Verhaltenskodizes gem. Art. 95 KI-VO. Unterschiede bestehen insbesondere bei der Bindungswirkung der Instrumente.[627] Verhaltenskodizes betreffen die freiwillige Selbstverpflichtung, Maßnahmen einzuhalten, die rechtlich nicht zwingend sind.[628] Der Selbstverpflichtung der Verhaltenskodizes steht die Nichtverbindlichkeit von Praxisleitfäden gegenüber, die als Orientierungshilfen dienen sollen.[629] Leitlinien hingegen sind weder reine Informationen noch mit Rechtsnormqualität ausgestattet, sie sind ein sog. unbenanntes Instrument, mangels Aufführung im Rechtsaktkatalog des Art. 288 AEUV.[630] Sie gewährleisten die Transparenz, Vorhersehbarkeit sowie Rechtssicherheit des Vorgehens der Kommission, indem diese sich an ihre Verwaltungspraxis bindet und bedingen in der Folge Vertrauensschutz für den Rechtsverkehr, von dem nicht behördenseitig willkürlich abgewichen werden darf.[631] Sie führen, vergleichbar norminterpretierenden und -konkretisierenden

[624] *Griesel/Brorsen*, MMR 2026, 477, 480.
[625] Nach Art. 96 Abs. 1 UAbs. 1 lit. d KI-VO sollen z. B. Leitlinien für die praktische Umsetzung der Transparenzpflichten des Art. 50 KI-VO erarbeitet werden.
[626] Praxisleitfäden werden auch in Art. 50 Abs. 7 KI-VO genannt, zur Erleichterung der wirksamen Umsetzung von Pflichten des Art. 50 KI-VO.
[627] BeckOK KI-Recht/*Schneider*, 7. Ed. 25.7.2026, KI-VO Art. 56 Rn. 7; BeckOK KI-Recht/*Brandt-Steinke*, 7. Ed. 1.8.2026, KI-VO Art. 96 Rn. 15f.
[628] BeckOK KI-Recht/*Gorzala*, 7. Ed. 1.8.2025, KI-VO Art. 95 Rn. 1f.; BeckOK KI-Recht/*Schneider*, 7. Ed. 25.7.2026, KI-VO Art. 56 Rn. 7.
[629] BeckOK KI-Recht/*Schneider*, 7. Ed. 25.7.2026, KI-VO Art. 56 Rn. 7.
[630] Martini/Wendehorst/*Hartmann*, 2. Aufl. 2026, KI-VO Art. 96 Rn. 1; BeckOK KI-Recht/*Brandt-Steinke*, 7. Ed. 1.8.2026, KI-VO Art. 96 Rn. 15.
[631] BeckOK KI-Recht/*Brandt-Steinke*, 7. Ed. 1.8.2026, KI-VO Art. 96 Rn. 15.

Verwaltungsvorschriften nach nationalem Recht, zu einer faktischen Bindungswirkung.[632] Überdies sind sie nach Art. 99 Abs. 1 KI-VO bei dem Erlass von Durchsetzungsmaßnahmen, wie Sanktionen, zu berücksichtigen.[633] Die Konkretisierungsinstrumente wurden auch bereits teilweise praktisch umgesetzt.[634]
Mit Blick auf Art. 50 KI-VO – aber auch die weiteren Pflichten in den folgenden Normen – besteht als Handlungsoption, **Leitlinien, Verhaltenskodizes** nach Art. 95 KI-VO sowie **Praxisleitfäden** nach Art. 56 KI-VO mit Blick auf die in solchen enthaltenen Ausführungen zur Transparenz und Information (und freilich auch mit Blick auf weitere Aspekte) zu verfolgen und regelmäßig auf Maßnahmen zu überprüfen, die im Rahmen des Einsatzes informationstechnischer Systeme bei der richterlichen Entscheidungsfindung vorteilhaft sind.[635] In solchen Praxisleitfäden sollen u. a. die Ziele eindeutig niedergelegt sein, die sich freilich prominent schon aus Art. 53 und 55 KI-VO ergeben, für die sie als „Soft Law“ – obgleich ihrer Unverbindlichkeit (einstweilen) – den erleichterten Nachweis der Rechtskonformität ermöglichen, vgl. Art. 53 Abs. 4 KI-VO und Art. 55 Abs. 2 KI-VO (sofern eine harmonisierte Norm nach Art. 40 KI-VO noch nicht in Kraft getreten ist).[636] Auch sollen sie die Erfüllung der Transparenzpflichten nach Art. 50 KI-VO unterstützen.[637] Die Leitfäden sollen Verpflichtungen und Maßnahmen enthalten, zur Gewährleistung der Verwirklichung dieser Ziele, die in konkrete Maßnahmenkataloge, bestenfalls Checklisten gefügt sind.[638] Hinsichtlich des Nachweises der Rechtskonformität kann durch die Befolgung der entsprechenden Leitfäden eine Verkürzung von Überprüfungen der Kommission, ob die Pflichten erfüllt werden, bedingt werden, sowie die Nachweisführung vereinfacht.[639]
Als weitere Handlungsoption kann zudem in den Blick genommen werden, in einen mitgliedstaatsübergreifenden Austausch über die zu implementierenden (Mindeststandard-)Maßnahmen und Best Practices zu gehen, um einen Erkenntnisgewinn durch Erfahrungsaustausch zu erhalten.[640] Auch kann ein – unional verständigter – freiwilliger Verhaltenskodex im Sinne des Art. 95 KI-VO erstellt werden, der über die in Kapitel III Abschnitt 2 der KI-VO genannten Anforderungen gem. Art. 95 Abs. 1 KI-VO hinaus, auch die Adressierung der Aspekte von Art. 50ff. KI-VO zum Gegenstand hat. Ein solches Vorgehen würde mit den Vorteilen derartiger

[632] BeckOK KI-Recht/*Brandt-Steinke*, 7. Ed. 1.8.2026, KI-VO Art. 96 Rn. 16.
[633] Siehe hierzu auch Martini/Wendehorst/*Hartmann*, 2. Aufl. 2026, KI-VO Art. 96 Rn. 2.
[634] So existieren bereits u. a. ein Verhaltenskodex für die Transparenz von KI-generierten Inhalten sowie Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme nach Artikel 50 der KI-Verordnung: *Bontcheva/Bechmann et al.,* Verhaltenskodex für die Transparenz von KI-generierten Inhalten „Code of Practice on Transparency of AI-Generated Content“, abrufbar unter https://digital-strategy.ec.europa.eu/de/policies/code-practice-ai-generated-content (Abruf am 03.09.2026, 15:58 Uhr); *Europäische Kommission,* Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme nach Artikel 50 der KI-Verordnung „Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the ‘AI Act’)“, abrufbar unter https://digital-strategy.ec.europa.eu/de/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems (Abruf am 31.08.2026, 20:51 Uhr).
[635] Siehe auch *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 27.
[636] Bomhard/Pieper/Wende/*Beurskens*, 1. Aufl. 2025, KI-VO Art. 56 Rn. 3, 19; BeckOK KI-Recht/*Schneider*, 7. Ed. 1.8.2026, KI-VO Art. 53 Rn. 34.
[637] BeckOK KI-Recht/*Schneider*, 7. Ed. 25.7.2026, KI-VO Art. 56 Rn. 4.
[638] Bomhard/Pieper/Wende/*Beurskens*, 1. Aufl. 2025, KI-VO Art. 56 Rn. 20.
[639] Martini/Wendehorst/*Bernsteiner/Schmitt*, 2. Aufl. 2026, KI-VO Art. 53 Rn. 62.
[640] Vgl. auch *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 21f., 27.

Verhaltenskodizes in Einklang stehen, die spezifischen Anforderungen bestimmter Branchen bzw. Themen genau – hier der richterlichen – adressieren zu können und unter dynamischer Weiterentwicklungsmöglichkeit eine (auch unionale) Verständigung, Harmonisierung und Standardisierung zu ermöglichen sowie regulatorische Lücken zu mildern.[641] Dem ähnelt das sinnvolle nationale Vorgehen nach der KI-Strategie der Justiz, wonach eine Bund/Länder-übergreifende Auslegungspraxis der KI-VO erarbeitet werden soll sowie weitere einheitliche Vorgehensweisen, insbesondere hinsichtlich Prozessen, Standards, Best Practices und Leitlinien verfolgt werden sollen.[642]
Des Weiteren kommt in Betracht, die Anforderungen der Art. 53 ff. KI-VO – dem Gedanken von Verhaltenskodizes folgend – unabhängig von der konkreten Ausgestaltung des KI-Modells zu implementieren, ob ein KI-Modell mit allgemeinem Verwendungszweck (Art. 3 Nr. 63 KI-VO) oder ggf. mit systemischem Risiko (Art. 3 Nr. 63 i. V. m. Art. 51 KI-VO) dem konkreten KI-System zugrunde liegt. Anbieter von KI-Modellen mit allgemeinem Verwendungszweck müssen die (Dokumentations- und Informations-)Pflichten nach Art. 53 KI-VO erfüllen, Anbieter von KI-Modellen mit allgemeinem Verwendungszweck mit systemischem Risiko hingegen des Weiteren die Pflichten nach Art. 55 KI-VO, die u. a. Modellbewertungspflichten sowie Pflichten zur Risikobewertung und -minderung umfassen.[643] Durch eine stete Implementierung kann einer ggf. diffizilen Abgrenzungsnotwendigkeit einzelner Modellarten zwischen den einzelnen Kategorien vorgebeugt werden und mit der freiwilligen Einhaltung überschießender Standards akzeptanzfördernd[644] geworben werden; auch sieht man sich hierdurch nicht der Gefahr ausgesetzt, versehentlich Pflichten nicht einzuhalten, die eigentlich befolgt werden müssten. KI-Modelle (nur) mit allgemeinem Verwendungszweck werden nämlich von solchen „mit systemischem Risiko“ anhand Art. 51 KI-VO abgegrenzt. Besitzen die Modelle (Art. 51 Abs. 1 lit. a KI-VO) „Fähigkeiten mit hohem Wirkungsgrad, die mithilfe geeigneter technischer Instrumente und Methoden, einschließlich Indikatoren und Benchmarks, bewertet werden“ oder verfügt ein Modell (Art. 51 Abs. 1 lit. b KI-VO) aufgrund einer Kommissionsentscheidung über eine Wirkung oder über Fähigkeiten, die denen nach lit. a entsprechen, sind diese KI-Modelle als solche mit allgemeinem Verwendungszweck mit systemischem Risiko einzustufen.[645]

bb. Innere Transparenz gegenüber Richterschaft und Betreiber

Die innere Transparenz betrifft die Nachvollziehbarkeit der Ergebnisse für die das System nutzenden Richter und Richterinnen sowie für den Betreiber. Sie ist Voraussetzung einer wirksamen menschlichen Aufsicht und Kontrolle (dazu bereits Abschnitt C.II.2) und einer tragfähigen Begründung der Entscheidung. Für Hochrisiko-KI-Systeme wirkt Art. 13 KI-VO konkretisierend.

[641] BeckOK KI-Recht/*Gorzala*, 7. Ed. 1.8.2025, KI-VO Art. 95 Rn. 15ff.
[642] *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 20, 25f., 28.
[643] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 290f.
[644] Vgl. zu den Verhaltenskodizes nach Art. 95 KI-VO BeckOK KI-Recht/*Gorzala*, 7. Ed. 1.8.2025, KI-VO Art. 95 Rn. 15.
[645] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 290.

(1) Umgang mit der sog. Black-Box-Problematik

Die Black-Box-Problematik, d. h. die fehlende Erklärbarkeit informationstechnischer Systeme ist primär bei Systemen maschinellen Lernens relevant.[646] Regelbasierte Ausgaben sind hiervon ausgenommen, da ihre Ausgabeergebnisse auf vordeterminierten Strukturen beruhen und stets replizierbar sowie damit nachvollziehbar sind, auch dann, wenn sie komplexen Logiken folgen mögen.[647] Transparenz staatlichen Handelns wird bei der richterlichen Entscheidungsfindung insbesondere im Nachgang, d. h. über die Urteilsbegründung hergestellt, in welcher die die Entscheidung tragenden Erwägungen in einer Weise dargelegt werden, die der Fremdkontrolle zugänglich ist. Die Begründung juristischer Entscheidungen hat eine Vielzahl von Funktionen. Sie dient der Selbstkontrolle, der Fremdkontrolle sowie der Generierung sozialer Akzeptanz und von Systemvertrauen.[648] Hinsichtlich des letzten Gesichtspunkts gilt, dass je genauer die Gründe dargelegt werden und je genauer die vorgebrachten Argumente verwertet werden, desto einfacher ist die Akzeptanz des Betroffenen gegenüber der Entscheidung zu erreichen, was generalisiertes Institutionsvertrauen fördern kann. Ein Fehlen einer Begründung kann indes Widerstand fördern und das Vertrauen in die Institution gefährden, weshalb Transparenzpflichten als vertrauensfördernde Maßnahmen auch mit Blick auf ein Vertrauen in die Technik in den Blick genommen werden sollten.[649]
Die Anforderungen und roten Linien zur Transparenzpflicht (insbesondere mit Blick auf die Entscheidungsbegründung) sind eingehend im Abschlussbericht des ersten Meilensteins dieses Projektes erarbeitet worden.[650] Systeme maschinellen Lernens können aufgrund ihrer Limitationen derzeit nicht entscheidungs*ersetzend* eingesetzt werden. Ein entscheidungs*unterstützender* Einsatz erfordert aufgrund der fehlenden Richtigkeitsgewähr eine sorgfältige Kontrolle und ggf. Korrektur der Arbeitsergebnisse dieser Systeme durch über die Funktionsweise dieser Systeme informierte Richterinnen und Richter.[651] Es ist hier sicherzustellen, dass die Entscheidung bzw. die Entscheidungsgründe, die KI-unterstützt erstellt worden sind, auch die tatsächlichen, die Entscheidung tragenden Gründe wiedergeben.
Für maßgebliche **Handlungsempfehlungen,** die eine informierte Kontrolle durch Richterinnen und Richter sicherstellen und zu den Empfehlungen der Information der Richterschaft über die konkrete Funktionsweise der eingesetzten KI-Systeme kann auf die Ausführungen im Abschnitt C.II.2 verwiesen werden, da sie dort bereits eingehend unter verschiedenen Aspekten der Aufsicht und Kontrolle erörtert worden sind.
Ergänzend lassen sich zur Unterstützung dieser Kontrollmechanismen mit Blick auf die Darstellung der Systemergebnisse mehrere Optionen andenken, wie die Nachvollziehbarkeit der Systemergebnisse für Richterinnen und Richter mit Blick auf die erforderliche Entscheidungsbegründung verbessert werden kann und sich so akzeptanzbildende Vertrauenstatbestände schaffen lassen: Hier kommt mit Blick auf die Nachvollziehbarkeit der Entscheidung des KI-Systems bereits eine systemseitige

---

[646] Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 5.
[647] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 149.
[648] *Wischmeyer*, AöR 143 (2018), 1, 56ff.
[649] *Wischmeyer*, AöR 143 (2018), 1, 58.
[650] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 146ff.
[651] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 155.

Strukturvorgabe der Systemausgabe in Betracht (z. B. mittels Strukturprompt), aufgrund derer das System bereits selbst seine (wahren) tragenden Gründe, adressatengerecht aufbereitet zur Verfügung stellt,[652] etwa indem es im Rahmen der Auswahl eines für den Richter bzw. die Richterin zur Verfügung gestellten Arguments oder einer zur Verfügung gestellten Rechtsansicht bereits als Textblock eine entsprechende Begründung in den Urteilstext überträgt oder sie im Rahmen einer Interaktion jedenfalls anzeigen lässt. Auch sollte das System die seiner Entscheidung entgegenstehenden Gründe sowie deren Entkräftung darstellen. So wäre mit den sonstigen Überprüfungs- und Kontrollinstrumenten eine menschliche Überprüfung zumindest der Argumentationsstruktur und der der Argumentation unterlegten Logik schnell möglich und auch dem Akzeptanzgesichtspunkt Rechnung getragen.[653]
Im Bereich der **regelbasierten Systeme** gilt auch für den dort möglichen Bereich des entscheidungsersetzenden Einsatzes: Die Darstellung der notwendigen Begründung lässt sich bei regelbasierten Systemen derart darstellen, dass im Rahmen der Entscheidungsbaumstruktur die Systemausgaben entsprechend gestaltet werden.[654] Die nutzerseitig (von den Parteien) eingegebenen Daten (d. h. die Eingaben) müssen mitsamt einer vorgefertigten Argumentation (vorformulierte Textbausteine) in einem von dem System ausgespielten Arbeitsergebnis (der Ausgabe) in einer logischen Argumentationsstruktur wiedergegeben werden, entsprechend der Tätigkeit einer menschlichen Richterin oder eines Richters.[655] Auf diese Weise werden die Transparenzerfordernisse gewahrt.

(2) Transparenzanforderungen an Hochrisiko-KI-Systeme nach Art. 13 KI-VO, insbesondere Betriebsanleitung

Sofern es sich bei KI-Systemen um Hochrisiko-KI-Systeme handelt, stellt die KI-Verordnung erhebliche und weitreichende Anforderungen an deren Gestaltung, Dokumentation, Protokollierung und Überwachung, vgl. Art. 9 ff. KI-VO. Sowohl bei entscheidungs*ersetzenden* als auch bei entscheidungs*unterstützenden* Systemen kommt eine solche Hochrisiko-Klassifizierung in Betracht, vgl. zu den Voraussetzungen Art. 6 KI-VO.[656] Liegt eine solche Klassifizierung vor, sind insbesondere die Transparenzerfordernisse nach Art. 13 KI-VO[657] (Transparenz und Bereitstellung von Informationen für die Betreiber) zu beachten.[658] Art. 13 Abs. 1 KI-VO stellt die funktionale Nachvollziehbarkeit des KI-Systems in den Vordergrund und Abs. 2 und 3 behandeln die Bereitstellung von Informationen über die Funktionsweise des KI-Systems in Form von Betriebsanleitungen.[659] Adressat der Verpflichtungen ist in erster Linie der Anbieter des Hochrisiko-KI-Systems.[660] Ratio der Bestimmung ist es, über transparente Systemgestaltung und Information den Betreiber in die Lage zu versetzen, das KI-System mit Verständnis und Beherrschbarkeit zu betreiben, vgl.

[652] *Wischmeyer*, AöR 143 (2018), 1, 61.
[653] *Quack*, Automation zivilrechtlicher Entscheidungsfindung, 1. Aufl. 2025, S. 149.
[654] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 150.
[655] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 151.
[656] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 282ff., 287ff., 291ff.
[657] Art. 13 KI-VO als Kern der Transparenzvorschriften bezeichnend: Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 11 Rn. 7.
[658] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 282ff., 287ff., 291ff.
[659] Hilgendorf/Roth-Isigkeit KI-VO/*Kumkar*, 2. Aufl. 2025, § 6 Rn. 16.
[660] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 13 Rn. 3; zustimmend BeckOK KI-Recht/*Zech/Linhart*, 7. Ed. 1.8.2026, KI-VO Art. 13 Rn. 16.

insoweit auch ErwGr. 72.[661] Transparenz soll hier durch Erklärbarkeit und Bereitstellung von Informationen erreicht werden.[662] Der Ansatz, den Art. 13 KI-VO verfolgt, ist ein System funktionsgerechter, abgestufter Transparenz, das Transparenzpflichten und Geheimhaltungspflichten angemessen balanciert.[663] Eine Quellcode-Veröffentlichung fordert Art. 13 Abs. 1 KI-VO damit auch nicht.[664] Die Transparenzerfordernisse entsprechen daher konzeptionell und inhaltlich wesentlich jenen des nationalen Rechts (dazu vorstehend). Explizit gefordert ist jedoch das Bereitstellen einer Betriebsanleitung für Betreiber, Abs. 2 und 3. Betriebsanleitungen sind im Produktsicherheitsrecht längst ein etabliertes Instrument der Kundeninformation,[665] weshalb hier – in Bezug auf die Gestaltungsmöglichkeiten – nach Orientierung gesucht werden könnte.

Art. 13 Abs. 1 KI-VO verlangt konkret, dass Hochrisiko-KI-Systeme so konzipiert und entwickelt werden, dass ihr Betrieb hinreichend transparent ist, damit die Betreiber die Ausgaben eines Systems angemessen interpretieren und verwenden können. Die Transparenz muss, so Art. 13 Abs. 1 S. 2 KI-VO, auf eine geeignete Art und in einem angemessenen Maß gewährleistet werden, damit die Anbieter und Betreiber ihre in Abschnitt 3 festgelegten einschlägigen Pflichten erfüllen können. Für Betreiber muss die Funktionsweise des Systems daher verständlich dargestellt werden, wobei es hier explizit nicht auf die vollständige Erklärbarkeit des Systems ankommt,[666] wie sich aus dem Wortlaut „*hinreichend transparent*“ ergibt, sondern lediglich darum, Betreibern die Funktionsweise, die Merkmale, Fähigkeiten aber auch die Leistungsbeschränkungen des KI-Systems zu verdeutlichen, vgl. so explizit ErwGr. 72. Somit sind auch Black-Box-Systeme nicht von der Verwendung ausgeschlossen, wenn nur hinreichend über deren Funktionsweise informiert wird und die Grenzen dieser Systeme aufgezeigt werden.[667]

Unter Betriebsanleitungen i. S. d. Art. 13 Abs. 2, 3 KI-VO versteht die KI-VO Informationen des Anbieters, die den Betreiber insbesondere über die Zweckbestimmung und die ordnungsgemäße Nutzung des KI-Systems informieren, Art. 3 Nr. 15 KI-VO. Die Betriebsanleitungen müssen „präzise, vollständige, korrekte und eindeutige Informationen in einer für den Betreiber relevanten, barrierefrei zugänglichen und verständlichen Form enthalten“. An die Betriebsanleitung werden somit inhaltliche und formelle Anforderungen gestellt. Der Inhalt der Betriebsanleitung steht in einem Zweckverhältnis zu dem in Art. 13 Abs. 1 KI-VO geregelten Transparenzgebot, sodass auch die Betriebsanleitung der Transparenz dient, wie ErwGr. 72 der KI-VO klarstellt.

Was die mindestens in der Betriebsanleitung **bereitzustellenden Informationen** sind, präzisiert Art. 13 Abs. 3 KI-VO. Hierzu zählen zum Beispiel die Kontaktdaten des Anbieters (lit. a), Angaben über Merkmale, Fähigkeiten und Leistungsgrenzen des Hochrisiko-KI-Systems (lit. b), sowie etwaige Änderungen (lit. c), Maßnahmen zur Gewährleistung der menschlichen Aufsicht gem. Art. 14 KI-VO (lit. d), Informationen über die erwartete Lebensdauer, erforderliche Wartungs- und Pflegemaßnahmen (lit. e) sowie gegebenenfalls eine Beschreibung der in das Hochrisiko-KI-System

[661] BeckOK KI-Recht/*Zech/Linhart*, 7. Ed. 1.8.2026, KI-VO Art. 13 Rn. 3; Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 13 Rn. 2.
[662] Zu den Zielrichtungen der Transparenz in der KI-VO Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 12 Rn. 4; von Wirkdimensionen der Transparenz sprechend Hilgendorf/Roth-Isigkeit KI-VO/*Kumkar*, 2. Aufl. 2025, § 6 Rn. 4ff.
[663] BeckOK KI-Recht/*Zech/Linhart*, 7. Ed. 1.8.2026, KI-VO Art. 13 Rn. 22.
[664] BeckOK KI-Recht/*Zech/Linhart*, 7. Ed. 1.8.2026, KI-VO Art. 13 Rn. 32.
[665] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 13 Rn. 2.
[666] BeckOK KI-Recht/*Zech/Linhart*, 7. Ed. 1.8.2026, KI-VO Art. 13 Rn. 2, 17.
[667] BeckOK KI-Recht/*Zech/Linhart*, 7. Ed. 1.8.2026, KI-VO Art. 13 Rn. 20f.

integrierten Mechanismen zur Erfassung, Speicherung und Auswertung von Protokollen gem. Art. 12 KI-VO (lit. f).[668] Begrenzt wird der Umfang der in der Betriebsanleitung transparent zu machenden Informationen durch berechtigte Geheimhaltungsinteressen des Anbieters.[669] Die formellen Anforderungen sind erfüllt, wenn die Informationen barrierefrei zugänglich und in einem geeigneten digitalen Format bereitgestellt oder auf andere Weise mit Betriebsanleitungen versehen sind.[670] Der Anbieter hat jedoch ein Wahlrecht, auch eine physische Betriebsanleitung vorzuhalten.[671] Die Barrierefreiheit liegt vor, wenn die Betriebsanleitungen in einer Weise zur Verfügung stehen, die Anbietern und Betreibern den Zugang unkompliziert ermöglicht.[672] Sie müssen daher in einfacher Sprache verfasst sein und auch eine logische Struktur aufweisen, die es Lesern mit körperlichen oder kognitiven Einschränkungen ermöglicht, Informationen schnell zu erfassen.[673] Bei der Gestaltung kann daher auf unterschiedliche Elemente wie visuelle Elemente, Audio- und Videoanleitungen, Brailleschrift, Videos in Gebärdensprache oder interaktive Gestaltungen zurückgegriffen werden.[674] Es bietet sich damit konkret an, eine den Anforderungen des Art. 13 Abs. 2, 3 KI-VO entsprechende Betriebsanleitung für das Hochrisiko-KI-System zu erstellen, vorzuhalten und diese den Betreibern bereitzustellen. Anbieten würde sich das digitale Format, wobei der Anbieter bei Auswahl der konkreten Gestaltungselemente frei ist, sofern er die normativen Zielvorgaben des Art. 13 Abs. 2 KI-VO einhält und die Mindestinhalte des Art. 13 Abs. 3 KI-VO aufnimmt. Eine solche Betriebsanleitung muss beim Test eines Hochrisiko-Systems außerhalb eines KI-Reallabors einem (zukünftigen) Betreiber zur Verfügung gestellt werden, vgl. Art. 60 Abs. 4 lit. h KI-VO. Sinn und Zweck dahinter ist, dass im Testbetrieb ebenfalls über die Zweckbestimmung und ordnungsgemäße Nutzung des KI-Systems informiert werden soll.
Als Handlungsempfehlung kann grundsätzlich auf die bereits erfolgten Empfehlungen zur menschlichen Aufsicht und Kontrolle verwiesen werden. Es sollte abstrakt über die Funktionsweise und das geplante Einsatzszenario des KI-Systems informiert werden. Weiterhin muss eine den Erfordernissen des Art. 13 Abs. 2, 3 KI-VO entsprechende Betriebsanleitung erstellt und bereitgestellt werden. Um Medienbrüche zu vermeiden, bietet sich hier ein digitales Format an, welches am besten in die Oberflächengestaltung des Systems aufgenommen werden kann und dort abrufbar ist.

cc. Überprüfbarkeit des Systems und Dokumentation

Über die Information der Beteiligten hinaus stellt sich die Frage, ob und inwiefern das System selbst einer vertieften Überprüfung zugänglich sein und seine Vorgänge dokumentieren muss.
Auch vor dem Hintergrund der Transparenz der systemseitigen Ergebnisfindung, Nachvollziehbarkeit der involvierten Logik sowie der wesentlichen Gründe der Entscheidungsfindung für die Parteien stellt ein **mehrstufiges Informationsmodell** eine wesentliche Handlungsoption dar, sodass verschiedene Informationstiefen, etwa

[668] Zu den Inhalten im Detail Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 13 Rn. 88ff.
[669] BeckOK KI-Recht/*Zech/Linhart*, 7. Ed. 1.8.2026, KI-VO Art. 13 Rn. 28; ebenfalls Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 13 Rn. 77.
[670] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 13 Rn. 83ff.
[671] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 13 Rn. 87.
[672] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 13 Rn. 85.
[673] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 13 Rn. 85.
[674] Martini/Wendehorst/*Martini*, 2. Aufl. 2026, KI-VO Art. 13 Rn. 86.

für Laien und für Experten, adressatengerecht bereitgestellt werden können.[675] Technische Komplexität lässt sich jedoch nicht verlustfrei reduzieren, ohne die Effektivität von Kontrollmöglichkeit zu beschränken; für eine erfolgreiche Missstandsidentifikation und -korrektur sind detaillierte Informationen zu den Systemen sowie Entscheidungsprozessen vorzuhalten.[676] Im Ergebnis muss zur tiefen und effektiven Kontrolle de facto Spezialisten Zugang zu dem System einschließlich der Trainingsmodelle, Datenbanken und Algorithmen etc. gegeben werden.[677]

(1) Quellcode-Offenlegung

Nach unserem Dafürhalten verlangt das geltende Recht keine Offenlegung des Quellcodes informationstechnischer Systeme für einen rechtskonformen Einsatz von KI im Justizsystem. Auch menschliche Entscheidungen sind nicht vollständig transparent, mit Blick auf einen auf Vertrauen basierenden Rechtsstaat stellt sich jedoch die Frage, ob nicht eine Quellcode-Veröffentlichung vertrauensbildende Effekte hätte.[678]
Eine **Handlungsoption** ist daher die proaktive, **vollständige (oder teilweise) Veröffentlichung des Quellcodes** in Verbindung mit einer Aufklärungspflicht durch die Entwickler sowie der zugrundeliegenden Trainingsdaten, damit entsprechend befähigte Parteien sowie Angehörige der Öffentlichkeit die Möglichkeit haben, die Aspekte der Entscheidungsfindung soweit als möglich nachzuvollziehen.[679] Dieser Blick in den justitiellen „Maschinenraum“ kann durch die einhergehende Transparenzerhöhung nicht nur der Akzeptanz von Rechtsstaat und Justiz im Besonderen, sondern zugleich auch der Akzeptanz des Einsatzes informationstechnischer Systeme bei der richterlichen Entscheidungsfindung zugutekommen, sowie das Auffinden systematischer Fehler durch eine breite, gesamtgesellschaftliche Kontrollmöglichkeit erleichtern.[680] Hierbei ist freilich auch das Thema der Geschäftsgeheimnisse, insbesondere von privaten Kooperationspartnern, bezüglich des Quellcodes zu beachten, sowie damit möglicherweise einhergehende Faktoren der Innovationshemmung.[681] Dem kann aber durch eine entsprechende Gestaltung der Kooperationsgrundlage begegnet werden, indem sich die öffentliche Hand (insbesondere bei erheblichem Einsatz öffentlicher Gelder) die Rechte am Quellcode bzw. die Veröffentlichungsmöglichkeit ausbedingt.[682] Im Gegenzug können beispielsweise längerfristige Kooperationsverträge angeboten werden.[683] Es sollte bei

[675] Vgl. *Wischmeyer*, AöR 143 (2018), 1, 46ff.
[676] *Wischmeyer*, AöR 143 (2018), 1, 53.
[677] *Wischmeyer*, AöR 143 (2018), 1, 53.
[678] Vgl. *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 337f.
[679] *Graichen*, Die Automatisierung der Justiz, 1. Aufl. 2022, S. 269, 280ff.; *Horn*, ZZP 136 (2023), 459, 491.
[680] *Horn*, ZZP 136 (2023), 459, 491.
[681] *Horn*, ZZP 136 (2023), 459, 491; Martini, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz, 1. Aufl. 2019, S. 182.
[682] Vgl. *Martini*, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz, 1. Aufl. 2019, S. 342; s. a. BeckOK DatenschutzR/*Schmidt-Wudy*, 57. Ed. 1.8.2026, DS-GVO Art. 15 Rn. 79, der den Praxistipp im Rahmen der Veröffentlichung von Scores gibt, sich vertragliche Garantien durch Dienstleister einräumen zu lassen, damit der Verantwortliche den Pflichten aus Art. 15 Abs. 1 lit. h DSGVO vollständig nachkommen kann.
[683] Wobei der Aspekt der technologischen Unabhängigkeit im Auge behalten werden sollte, Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_z

einer vollständigen Veröffentlichung der Systemgrundlagen jedoch auch – neben der Tatsache, dass sensible Sicherheitsaspekte aufgedeckt werden (können) – beachtet werden, dass Bürgerinnen und Bürger regelmäßig nicht über derartige Fachkenntnisse verfügen und die für ihr Verständnis notwendigen Erkenntnisse verborgen bleiben können.[684] Zu berücksichtigen ist mit Blick auf Sicherheitsrisiken auch, dass v. a. KI-Systeme mit Coding-Funktionalität bereits erfolgreich eingesetzt worden sind, um Schwachstellen von staatlichen Systemen zu entdecken und zu umgehen (siehe oben Abschnitt C.II.3.b.aa). Dies wird vereinfacht durch einen offenen Umgang mit Quellcodes der staatlichen Systeme. Die **Handlungsoption** konfligiert daher mit Sicherheitsinteressen und **sollte nicht umgesetzt** werden.

Eine **weitere, gegenüber der voranstehenden zu befürwortende Handlungsoption** ist die **restriktive Handhabe der Quellcode-Veröffentlichung**. Die Offenlegung des Quellcodes kann ausnahmsweise notwendig sein, wenn das Offenlegungsinteresse hieran gegenläufige Interessen überwiegt.[685] Dies kann beispielsweise in einem Rechtsmittelverfahren relevant werden, wenn in dem konkreten Fall berechtigte Zweifel daran bestehen, dass ein entscheidungsersetzendes informationstechnisches System akkurat funktioniert hat.[686] In diesem Fall kann gutachterlich Beweis über die relevanten Aspekte erhoben werden, indem der Quellcode einem Gerichtsgutachter zugänglich gemacht wird. Sollte hingegen eine relevante Fehlfunktion bezüglich des informationstechnischen Systems bereits unabhängig vom konkreten Fall im Rahmen der (allgemeinen) Aufsicht und Kontrolle (siehe hierzu bereits Abschnitt C.II.2.a.bb(1)) hervortreten, kann die Fehlfunktion durch das jeweils vorgesehene Überwachungsgremium (intern) behoben werden, ohne die Notwendigkeit der Offenlegung des Quellcodes nach außen.

In den relevanten Fällen einer notwendigen Offenlegung sollte indes bedacht werden, dass durch diese u. a. Geschäftsgeheimnisse und Urheberrechte der Unternehmen betroffen sein können, die die Software entwickeln, sowie Sicherheitsaspekte im Hinblick auf das System selbst (z. B. mögliche Schwachstellen zur manipulativen, externen Einflussnahme).[687] Diese Ausnahmefälle der Offenlegung sollten daher bereits vorab bedacht und entsprechende Maßnahmenprotokolle ausgearbeitet werden. Hierbei sollten auch die (Belange der) entsprechenden Stakeholder (z. B. ein das System betreuendes Unternehmen) einbezogen werden, deren Interessen berührt werden könnten. Da derartige Offenlegungen erhebliche, schwerlich reversible Folgen nach sich ziehen können, sollten derartige Konstellationen, der Tragweite der Entscheidung entsprechend, von einem mit mehreren Richterinnen und Richtern besetzten Spruchkörper oder einer unabhängigen mit Fachkenntnis ausgestatteten, sensibilisierten Institution entschieden werden.[688] Dies wäre im Fall der Systemüberprüfung im Rechtsmittelkontext gem. § 75 GVG bzw. § 122 GVG bereits gegeben.[689]

Nach Möglichkeit sollte das System bei der Gestaltung auch technikseitig auf Offenlegungsfälle vorbereitet sein, sodass nur die notwendigen Informationen

---

ukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 127.

[684] *Nink*, Justiz und Algorithmen, 1. Aufl. 2021, S. 339.

[685] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 148.

[686] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 148.

[687] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 148.

[688] Vgl. auch zur Frage der Informationspflichterfüllung im Kontext von Geschäftsgeheimnissen BeckOK DatenschutzR/*Schmidt-Wudy*, 57. Ed. 1.8.2026, DS-GVO Art. 15 Rn. 78.3.

[689] Zu den Kammer- bzw. Senatsbesetzungen siehe etwa BeckOK GVG/*Feldmann*, 32. Ed. 15.8.2026, GVG § 75 Rn. 1ff.; BeckOK GVG/*Conrad-Graf*, 32. Ed. 15.8.2026, GVG § 122 Rn. 1ff.

herausgegeben werden. Die entsprechende Abstimmung zwischen den technisch und rechtlich für diese Fragen Zuständigen sollte eng erfolgen. Eine bereichsspezifische sowie modulare Entwicklung von informationstechnischen Systemen erscheint auch vor diesem Hintergrund vorteilhaft, da hiermit in den ausnahmsweisen Fällen, bei denen der Quellcode des informationstechnischen Systems offengelegt werden muss,[690] nicht der gesamte Code des gesamten justitiellen informationstechnischen Systems mit einem Mal offenbar wird, sondern nur ein betroffener Teilbereich.
Erwägbar erscheint auch (zusätzlich) ein präventives Kontrollregime, das mit entsprechend kompetenten Experten besetzt ist.[691] Eine Zertifizierung durch ein solches präventives Kontrollregime kann z. B. als zwingende Voraussetzung für den rechtmäßigen Einsatz des informationstechnischen Systems ausgestaltet sein (siehe zur Thematik der Zertifizierung auch Abschnitt C.II.3.b.dd).[692] Die Fälle der nachträglichen Überprüfung(smöglichkeiten) sollten hiervon unbenommen sein und die präventive sowie die Überprüfung im Nachhinein als sich ergänzendes Ganzes angesehen werden.[693]

(2) Protokollierungspflichten durch Anbieter und Betreiber

Transparenz kann auch durch Dokumentation erreicht werden; diese Wirkungsdimension der Transparenz ist in Art. 12 KI-VO verankert.[694] Die Dokumentation durch Protokollierung soll eine Rückverfolgbarkeit der konkreten Ereignisse der KI-Systeme ex post ermöglichen[695] und damit der sog. Black-Box-Problematik entgegenwirken.[696] Diese Aufzeichnungspflichten stellen dabei eine von mehreren Maßnahmen zur Beobachtung von KI-Systemen nach dem Inverkehrbringen dar und ergänzen damit die technische Dokumentation aus Art. 11 KI-VO.[697] Eine Protokollierung muss ausweislich des Wortlauts von Art. 12 KI-VO während des gesamten Lebenszyklus des Hochrisiko-KI-Systems stattfinden. Die Log-Files bzw. Protokolle müssen mindestens sechs Monate aufbewahrt werden, vgl. Art. 19 Abs. 1, 26 Abs. 6 UAbs. 1 KI-VO. Die genaue Aufbewahrungsdauer ist jedoch von der Zweckbestimmung des Hochrisiko-KI-Systems abhängig, sodass die Log-Files im Einzelfall unterschiedlich lange aufbewahrt werden müssen. Die Zweckbestimmung ist gem. Art. 3 Nr. 12 KI-VO „*die Verwendung, für die ein KI-System laut Anbieter bestimmt ist, einschließlich der besonderen Nutzungsumstände und Nutzungsbedingungen entsprechend den Angaben des Anbieters in der Gebrauchsanweisung, im Werbe- oder Verkaufsmaterial und in diesbezüglichen Erklärungen sowie in der technischen Dokumentation*“. Für die Nutzung bei der richterlichen Entscheidungsfindung sollte daher eine längere Aufbewahrungsdauer als die Mindestdauer gewählt werden. Die Aufbewahrungspflicht ist jedoch durch weitere rechtliche Verpflichtungen beschränkt.[698] So hat insbesondere das Datenschutzrecht, aber auch sonstiges Unionsrecht und nationales Recht Vorrang.[699] Die Grundsätze der

[690] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 148.
[691] *Wischmeyer*, AöR 143 (2018), 1, 63.
[692] *Wischmeyer*, AöR 143 (2018), 1, 63.
[693] Vgl. *Wischmeyer*, AöR 143 (2018), 1, 63.
[694] Hilgendorf/Roth-Isigkeit KI-VO/*Kumkar*, 2. Aufl. 2025, § 6 Rn. 5ff.
[695] Hilgendorf/Roth-Isigkeit KI-VO/*Kumkar*, 2. Aufl. 2025, § 6 Rn. 14.
[696] Reusch/Chibanguza/*Hessel*, 1. Aufl. 2026, KI-VO Art. 12 Rn. 2.
[697] Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 12 Rn. 5f.
[698] Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 12 Rn. 11.
[699] Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 12 Rn. 11.

Datenlöschung und Datensparsamkeit müssen berücksichtigt werden,[700] wie dieses Spannungsfeld jedoch konkret aufzulösen ist, ist unklar.[701] Ähnlich wie die Länge der Aufbewahrungspflicht bestimmt sich auch der angemessene Maßstab der Rückverfolgbarkeit der durch das Hochrisiko-KI-System getroffenen Entscheidungen nach der Zweckbestimmung, vgl. Art. 12 Abs. 2 KI-VO. Die Protokollierungspflicht ist zum einen auf die bestimmungsgemäße Verwendung des KI-Systems beschränkt und zum anderen ist der Einsatzbereich zu berücksichtigen.[702]
Die Aufbewahrungspflicht für automatische Protokolle gem. Art. 12 Abs. 1 KI-VO adressiert entweder den Anbieter nach Art. 19 Abs. 1 KI-VO oder den Betreiber nach Art. 26 Abs. 6 KI-VO. Wen die konkrete Aufbewahrungspflicht trifft, hängt davon ab, wer die Kontrolle über die Protokolle hat, vgl. Art. 19 Abs. 1 S. 1, 26 Abs. 6 UAbs. 1 KI-VO.[703] Diese parallele Regelung einer Aufbewahrungspflicht stellt sicher, dass die Aufbewahrungspflicht in jedem Fall eingehalten wird.[704] Die Kontrolle ist faktisch zu bestimmen.[705] Als Anbieter kommt die Ziviljustiz in Betracht;[706] Betreiber kann sie aber je nach Ausgestaltung des Betriebs und der Entwicklung des Systems ebenfalls sein.[707] Geht man beispielsweise durch den justizeigenen Entwicklungsvorgang in Kooperation mit einem Entwicklungsunternehmen davon aus, dass die Justiz sowohl Anbieter des Systems als auch durch den Betrieb im eigenen System auch Betreiber ist, trifft die Aufbewahrungspflicht in jedem Fall die Justiz als solche. Dies mag allerdings bereits divergieren, wenn ein in einem Bundesland entwickeltes System von anderen Ländern in deren eigener Domäne, d. h. im eigenen IT-Betrieb betrieben wird. So mag die entwickelnde Justiz zwar Anbieter sein, die Betreiberrolle ist jedoch dann nicht mehr gegeben. Die faktische Kontrolle der Protokolle liegt dann (je nach Gestaltung) in der Regel nicht mehr beim Anbieter, so dass die Aufbewahrungspflicht den die Kontrolle ausübenden Betreiber trifft.
Als **Handlungsempfehlung** ist somit lediglich allgemein zu geben, dass je nach Gestaltungssituation des Einsatzes des KI-Systems die technische Protokollierung beim Einsatz von KI im Rahmen der richterlichen Entscheidungsfindung und deren Aufbewahrung gemäß den vorstehend genannten Rahmenbedingungen sichergestellt wird. Der notwendige Maßstab der Rückverfolgbarkeit durch die Protokollierung sowie die konkrete Dauer der Aufbewahrungspflicht der Log-Files bestimmen sich hierbei nach der Zweckbestimmung, wie dies die KI-VO in Art. 12 Abs. 2, 19 Abs. 1, 26 Abs. 6 UAbs. 1 vorgibt. Orientierung können hier das Funktionscluster und die möglichen Einsatzszenarien[708] schaffen. Die Protokollierung sollte aufgrund des sensiblen Einsatzbereichs so genau sein, dass damit möglichst jede Entscheidung des KI-Systems erklärbar ist. Die Ziviljustiz als betroffener Adressat kann sich bei der konkreten Aufbewahrungsdauer an der

[700] *Spindler*, CR 2021, 361, 368.
[701] *Bomhard/Etzkorn*, ZdiW 2022, 159, 161.
[702] Martini/Wendehorst/*Braun Binder/Egli*, 2. Aufl. 2026, KI-VO Art. 12 Rn. 13.
[703] Die Sonderregelungen für Finanzinstitute spielen im hiesigen Kontext keine Rolle, zu ihnen aber Martini/Wendehorst/*I. Eisenberger/Th. Binder*, 2. Aufl. 2026, KI-VO Art. 19 Rn. 19ff. und Art. 26 Rn. 46.
[704] Martini/Wendehorst/*I. Eisenberger/Th. Binder*, 2. Aufl. 2026, KI-VO Art. 19 Rn. 3.
[705] Reusch/Chibanguza/*Heuer-James*, 1. Aufl. 2026, KI-VO Art. 19 Rn. 4; Martini/Wendehorst/*I. Eisenberger/Th. Binder*, 2. Aufl. 2026, KI-VO Art. 19 Rn. 17.
[706] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 278; ebenso *Wachter/Leeb*, RDi 2024, 440, 441.
[707] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 280; ebenso *Wachter/Leeb*, RDi 2024, 440, 441.
[708] Hierzu ausführlich *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 33ff.

Justizaktenaufbewahrungsverordnung[709] (JAktAV) orientieren, welche aufgrund des Justizaktenaufbewahrungsgesetzes[710] (JAktAG) erlassen wurde. Zivilakten ohne fortwirkenden Titel müssen fünf Jahre aufbewahrt werden. Akten, die zur Zwangsvollstreckung geeignete Titel und Beschlüsse enthalten, müssen 30 Jahre aufbewahrt werden. Der Fristbeginn bei der Aufbewahrungspflicht aus der KI-VO und derjenigen aus § 1 JAktAG ist jedoch unterschiedlich. Nach § 4 Abs. 1 JAktAV beginnt die Frist mit Ablauf des Jahres, in dem die Weglegung der Akten angeordnet wurde. Für Akten, für die die Weglegung nicht geregelt ist, beginnt die Aufbewahrungs- und Speicherungsfrist mit Ablauf des Jahres, in dem die letzte Verfügung in der Sache ergangen ist. Die Aufbewahrungspflicht der Protokolle aus der KI-VO beginnt mit dem Ende ihrer technischen Erzeugung. Daher kann nicht einfach der Fristenzeitraum übernommen werden, sondern die Aufbewahrungspflicht der Protokolle sollte gesetzlich mit dem Fristende aus der JAktAV synchronisiert werden.

c. Zusammenfassung

Der Einsatz informationstechnischer Systeme bei der richterlichen Entscheidungsfindung muss den Transparenzanforderungen staatlichen Handelns genügen. Die Parteien des Verfahrens und die Öffentlichkeit müssen über das „Ob“ und das „Wie“ des Einsatzes informationstechnischer Systeme bei der richterlichen Entscheidungsfindung informiert werden. Weiter ist die abstrakte Funktionalität des Systems transparent zu machen. Hinzu treten die Pflichten nach Art. 50 ff. KI-VO und bei Hochrisiko-KI-Systemen die gegebenenfalls weiterreichenden Anforderungen insbesondere der Art. 13, 14 Abs. 4, 26 und 27 KI-VO. Eine Quellcode-Offenlegung ist grundsätzlich nicht erforderlich. Die sogenannte Black-Box-Problematik bei Systemen maschinellen Lernens steht der Erfüllung einer Transparenz- und Informationspflicht nicht entgegen, da eine nachträgliche Entscheidungsbegründung dem Transparenzgebot hinreichend Rechnung trägt.

Die **Mitteilung des Einsatzes des informationstechnischen Systems** sollte vor bzw. spätestens zum Verfahrensbeginn erfolgen und muss die Information über die Art des Systems und seine abstrakte Funktionslogik beinhalten. Um den Anforderungen an die **äußere Transparenz** zu genügen und einer „Informationsüberflutung“ und einer „Desinformation durch Information“ vorzubeugen, besteht die Möglichkeit eines **abgestuften, verständlichen und barrierefreien Informationsmodells**, das möglichst **in die Verfahrenskommunikation eingebettet** ist. Die Informationen könnten zusätzlich insbesondere über Justizwebseiten, gerichtliche Schreiben oder in physischer Form bereitgestellt werden. Soweit Parteien im entscheidungsersetzenden Einsatz ausschließlich mit dem System interagieren, kann der Verfahrensfortgang von einer Bestätigung der Kenntnisnahme der zu verschaffenden Informationen abhängig gemacht werden. Dies wäre für die beklagte Partei an die Geltendmachung von Verteidigungsmitteln zu knüpfen, um ihr kein unberechtigtes Blockadepotential zukommen zu lassen. Zu beachten sind weiterhin die Pflichten aus Art. 26 Abs. 11 KI-VO und Art. 86 KI-VO sowie die Transparenzpflichten der Art. 50 ff. KI-VO.

Hinsichtlich der **Transparenzpflichten nach Art. 50 KI-VO** ist Folgendes festzuhalten: Die notwendigen Transparenzpflichten sollten in die prozessuale

[709] Justizaktenaufbewahrungsverordnung vom 8. November 2021 (BGBl. I S. 4834), die zuletzt durch Artikel 3 des Gesetzes vom 29. März 2026 (BGBl. 2026 I Nr. 83) geändert worden ist.
[710] Justizaktenaufbewahrungsgesetz vom 22. März 2005 (BGBl. I S. 837, 852), das durch Artikel 3 des Gesetzes vom 16. Oktober 2020 (BGBl. I S. 2187) geändert worden ist.

Verfahrenskommunikation bzw. Arbeitskommunikation direkt eingebettet werden. Im entscheidungsersetzenden Einsatz sollte die Interaktion mit einem KI-System stets klar offengelegt werden. Dies kann z. B. durch eine **gut sichtbare Kopfzeile** oder eine dem **KI-System vorgeschaltete Eingangssequenz** erfolgen. Im entscheidungsunterstützenden Einsatz kann die Interaktion mit einem KI-System gegenüber den Richterinnen und Richtern offengelegt werden, was aber regelmäßig nicht erforderlich ist. In diesem Einsatzszenario besteht gegenüber den Parteien hingegen mangels direkter Interaktion keine Offenlegungspflicht. Künstlich hergestellte Inhalte sind nach Art. 50 Abs. 2 KI-VO menschen- und maschinenlesbar als **KI-generierte Inhalte zu kennzeichnen**. Bei den technischen Lösungen hierfür ist der allgemein anerkannte Stand der Technik zu berücksichtigen. Bei der Umsetzung der Transparenzpflichten sind mögliche Kollisionen verschiedener Belange zu beachten (z. B. Datenschutzaspekte). Im **entscheidungsersetzenden Einsatz** kommt ein **textlicher Zusatz in der Entscheidung, dass diese automatisiert generiert worden ist**, in Betracht. In einer elektronisch abrufbaren Version kann z. B. ein maschinenlesbarer Hinweis ergänzt werden. Eine Kennzeichnungspflicht besteht dem Grunde nach auch im Fall des entscheidungsunterstützenden Einsatzes. Die Ausnahmen von der Kennzeichnungspflicht nach Art. 50 Abs. 2 S. 3 KI-VO können bei administrativen Funktionalitäten in Betracht kommen, sind jedoch für weitergehende sachlich-inhaltliche Funktionalitäten nicht einschlägig. Auch im richterlichen Kontext im entscheidungsunterstützenden Einsatz ist eine Kennzeichnung im Zweifel sinnvoll. Die Kennzeichnungspflicht des Anbieters nach Art. 50 Abs. 2 KI-VO geht nicht weiter als die Offenlegungspflicht des Betreibers nach Art. 50 Abs. 4 KI-VO, im Fall der Weiterverarbeitung „synthetisch“ generierten Texts durch Nutzerinnen oder Nutzer, beispielsweise bei der Generierung eines Textteils für ein Urteil durch Richterinnen oder Richter, wenn dieser im finalen Urteil konkret angepasst und weiterverarbeitet worden ist.

Die **Offenlegungspflicht der Betreiber** nach Art. 50 Abs. 4 UAbs. 2 KI-VO erfasst Gerichtsentscheidungen im Kontext des entscheidungsersetzenden sowie des entscheidungsunterstützenden Einsatzes informationstechnischer Systeme, nach derzeitigem Stand jedoch nicht gerichtliche Hinweise und Mitteilungen an die Verfahrensbeteiligten. Sie kann insbesondere durch **standardmäßige Hinweise an festgelegten Stellen oder Icon-Kennzeichnungen** ausgefüllt werden. Im Rahmen des entscheidungsersetzenden Einsatzes greift die Offenlegungspflicht nach Art. 50 Abs. 4 UAbs. 2 KI-VO, in der Regel ohne Durchgreifen der Ausnahmeregelung. Im Rahmen des **entscheidungsunterstützenden Einsatzes** besteht zwar die **Offenlegungspflicht** nach Art. 50 Abs. 4 UAbs. 2 KI-VO dem Grunde nach, **entfällt** allerdings **meist** ausnahmeregelungsbedingt **durch die richterliche Endkontrolle und Verantwortungsübernahme**. Mit Blick auf die von der KI-VO unberührten verfassungsrechtlichen Transparenzerfordernisse ist eine **Information über das „Ob“ und das „Wie“** des KI-Einsatzes **allerdings auch hier zu empfehlen**.

Zur Konkretisierung der Pflichten nach Art. 50 ff. KI-VO bietet es sich an, **Leitlinien, Praxisleitfäden und Verhaltenskodizes** zu verfolgen und regelmäßig auf vorteilhafte Maßnahmen zu überprüfen. Zudem kommen ein mitgliedstaatsübergreifender Austausch über die zu implementierenden **(Mindeststandard-)Maßnahmen und Best Practices**, ein unional verständigter Verhaltenskodex nach Art. 95 KI-VO für den richterlichen Bereich und die Einhaltung der Anforderungen der Art. 53 ff. KI-VO unabhängig von der konkreten Ausgestaltung des KI-Modells in Betracht.

Die **innere Transparenz gegenüber den nutzenden Richterinnen und Richtern** sowie dem Betreiber ist Voraussetzung wirksamer menschlicher Aufsicht und einer tragfähigen Entscheidungsbegründung. Die Black-Box-Problematik bei informationstechnischen Systemen maschinellen Lernens steht einem Einsatz nicht generell entgegen. Es ist sicherzustellen, dass die Entscheidungen bzw. die Entscheidungsgründe, die KI-unterstützt erstellt worden sind, auch die tatsächlichen, die Entscheidung tragenden Gründe wiedergeben. Ergänzend zu Empfehlungen hinsichtlich der menschlichen Aufsicht und Kontrolle kommt mit Blick auf die Nachvollziehbarkeit der Entscheidung des KI-Systems eine systemseitige Strukturvorgabe in Betracht, nach der das System seine tragenden Gründe adressatengerecht aufbereitet und dabei auch entgegenstehende Gründe samt deren Entkräftung darstellen sollte. Die Darstellung der notwendigen Begründung lässt sich bei regelbasierten Systemen derart darstellen, dass im Rahmen der Entscheidungsbaumstruktur die Systemausgaben entsprechend gestaltet werden, wodurch auch im entscheidungsersetzenden Einsatz die Transparenzerfordernisse gewahrt werden können.
Für Hochrisiko-KI-Systeme verlangt Art. 13 KI-VO keine vollständige Erklärbarkeit oder Veröffentlichung des Quellcodes. Gefordert wird eine hinreichende, funktionsgerechte Transparenz. Nach Art. 13 Abs. 2, 3 KI-VO ist eine **Betriebsanleitung** zu erstellen und bereitzustellen. Diese **sollte möglichst digital in die Systemoberfläche integriert werden**.
Für die Überprüfbarkeit des Systems ist ein mehrstufiges Informationsmodell eine wesentliche Handlungsoption. Eine tiefgehende Kontrolle setzt letztlich den Zugang von Spezialisten zum System und seinen Bestandteilen voraus.
Die **proaktive Veröffentlichung des Quellcodes**, die das geltende Recht nicht verlangt, könnte zwar vertrauensbildend wirken, gerät jedoch mit Sicherheitsinteressen in Konflikt und **sollte** daher **nicht umgesetzt werden**. Zu befürworten ist vielmehr eine restriktive Handhabung der Quellcode-Offenlegung. Die Offenlegung des Quellcodes kann ausnahmsweise notwendig sein, wenn das Offenlegungsinteresse hieran gegenläufige Interessen überwiegt, was z. B. in einem Rechtsmittelverfahren relevant werden kann, bei berechtigten Zweifeln an der akkuraten Funktion eines entscheidungsersetzenden Systems. Der Quellcode kann in diesem Fall einem Gerichtsgutachter zum Zwecke der Beweiserhebung zugänglich gemacht werden. Unabhängig davon auftretende Fehlfunktionen können dagegen intern im Rahmen der (allgemeinen) Aufsicht und Kontrolle behoben werden. Die Ausnahmefälle der Offenlegung sollten vorab bedacht und entsprechende Maßnahmenprotokolle ausgearbeitet werden. Diese Konstellationen sollen von einem mit mehreren Richterinnen und Richtern besetzten Spruchkörper oder einer unabhängigen sachkundigen Institution entschieden werden. Das System sollte auch technikseitig auf diese Fälle vorbereitet sein, sodass nur notwendige Informationen herausgegeben werden. Ergänzend ist ein präventives Kontrollregime erwägbar.
Schließlich ist die **technische Protokollierung** nach Art. 12 KI-VO sicherzustellen. Sie sollte so genau sein, dass möglichst jede Entscheidung des Systems erklärbar ist. Die Aufbewahrungspflicht trifft in der Regel die das System im eigenen IT-System betreibende Justiz. Die Aufbewahrungsdauer sollte mit den Fristen aus der JAktAV synchronisiert werden.

5. Datenschutz

a. Übergreifende rote Linien für jeden Einsatz

In einem zivilgerichtlichen Verfahren werden regelmäßig personenbezogene, mithin schutzwürdige Daten erhoben, verarbeitet und gespeichert, so beispielsweise bereits Name, Adresse und sonstige Informationen zu Kläger und Beklagtem.[711]
Aus dem **Datenschutzrecht** ergeben sich zahlreiche rote Linien für einen entscheidungs*ersetzenden* und *-unterstützenden* Einsatz informationstechnischer Systeme, die wir in dem Abschlussbericht unseres ersten Meilensteins dieses Projekts eingehend herausgearbeitet haben.[712] Relevant sind insoweit das Grundrecht auf informationelle Selbstbestimmung gemäß Art. 2 Abs. 1, 1 Abs. 1 GG[713], Art. 8 Abs. 1 EMRK[714] und Art. 8 Abs. 1 GRCh[715] sowie die Vorschriften der DSGVO[716].
Die roten Linien beziehen sich zunächst auf die **Verarbeitung potenziell entscheidungsrelevanter, personenbezogener Daten im Rahmen der Entscheidungsfindung** selbst. Für die Datenverarbeitung bedarf es einer entsprechenden rechtlichen Grundlage.[717] Die Verarbeitung personenbezogener Daten ist auch bei Einsatz informationstechnischer Systeme im Rahmen der Entscheidungsfindung grundsätzlich zulässig, da sie der Durchsetzung subjektiver Rechte, der Sicherstellung der Funktionsfähigkeit der Justiz und der Gewährung effektiven Rechtsschutzes dient und somit von einer legitimen Zwecken dienenden Rechtsgrundlage getragen ist, vgl. insbesondere Art. 8 Abs. 2 EMRK, Art. 6 Abs. 1 lit. e Alt. 2 DSGVO, Art. 9 Abs. 2 lit. f Alt. 2 DSGVO, Art. 23 Abs. 1 lit. f und j DSGVO.[718] Besonderheiten ergeben sich bei einem entscheidungs*ersetzenden* Einsatz deshalb, weil hierfür gem. Art. 22 DSGVO zunächst eine Rechtsgrundlage geschaffen werden müsste.[719]
Ein weiterer zu beachtender und wichtiger Aspekt des Datenschutzrechts betrifft die Gewährleistung von **Datensicherheit**. Werden personenbezogene Daten verarbeitet, sind diese Datenverarbeitungs- und Datenspeicherungsvorgänge so sicher auszugestalten, dass z. B. unbefugte Dritte keinen Zugriff auf diese Daten erlangen können. Datensicherheit ist daher bereits bei der Systemgestaltung hinreichend Rechnung zu tragen, wobei insbesondere die Art. 24, 25, 32, 35 DSGVO zu nennen sind.[720] Die Eingabe personenbezogener Daten in solche Systeme, die nicht der Kontrolle der Justiz unterstehen bzw. die datenschutzrechtlichen Anforderungen nicht

---

[711] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 96, 252ff., 320, 325, vgl. auch S. 153.
[712] Siehe hierzu ausführlich *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025.
[713] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 96ff.
[714] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 252ff.
[715] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 265ff.
[716] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 301ff. (s. a. insb. 319ff.)
[717] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 95f., 97ff., 101f., 252ff., 265ff., 307ff., 315ff., 320, 323ff., 331.
[718] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 96ff., 252ff., 265ff., 319ff.
[719] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 315ff., 323ff., 331.
[720] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 94ff., 260ff., 306.

wahren, z. B. ein cloudbasiertes Dialogsystem wie ChatGPT, ist in der Regel daher unzulässig.[721]
Auch mit Blick auf das **Training von KI-Systemen** ergeben sich aus dem Datenschutzrecht Anforderungen: Als Trainingsdaten sind v. a. nicht-personenbezogene Daten oder anonymisierte Daten zu verarbeiten, soweit dies möglich ist.[722] Soweit personenbezogene Daten verwendet werden, ist eine Verarbeitung zulässig, soweit bei einer Abwägung zwischen den von der Justiz verfolgten legitimen Zwecken (z. B. die Durchsetzung subjektiver Rechte, die Sicherstellung der Funktionsfähigkeit der Justiz und die Gewährung effektiven Rechtsschutzes) und den Betroffenenrechten erstere überwiegen. Insoweit ist eine Verwendung zulässig, wenn sie zum Zwecke der Entscheidungsfindung erforderlich ist. Grundsätzlich unzulässig ist sie hingegen, soweit sie nicht mehr unmittelbar für die konkret zu treffende Entscheidung relevant ist.[723] Werden personenbezogene Daten, die legitim im gerichtlichen Verfahren gem. § 3 BDSG erhoben worden sind, für das Training des informationstechnischen Systems weiterverwendet, bedarf es für diese Weiterverwendung, die einen neuen Eingriff darstellt, einer neuen Rechtfertigung. Ist ein Verzicht auf das Training mit personenbezogenen Daten nicht möglich, da dies für die Funktionsfähigkeit des Systems zwingend erforderlich ist, bedarf es einer entsprechenden Verhältnismäßigkeitsprüfung, die funktionsbezogen abwägt und die betroffenen Interessen in einen Ausgleich bringt.
Die **Löschung** der personenbezogenen Daten ist (gegebenenfalls durch Anonymisierung) ebenfalls zu ermöglichen, sobald der Anlass für die Datenerhebung entfallen ist oder das Löschungsinteresse überwiegt.[724] Auch dies ergibt sich aus den vorstehend genannten Verbürgungen als rote Linie. Ist die Umsetzung eines Löschungsrechts unmöglich, da die gespeicherten Daten sich nicht löschen lassen, ist ein Einsatz eines Systems ebenfalls unzulässig. Allerdings kommt eine Löschung dann nicht in Betracht, wenn die Funktionsfähigkeit des informationstechnischen Systems von der Speicherung der Daten abhängt oder diese Datenspeicherung gesetzlich vorgeschrieben ist.[725]
Zudem bestehen aufgrund datenschutzrechtlicher Verbürgungen **Informationspflichten** betreffend die Datenverarbeitung, die in allen Phasen des gerichtlichen Verfahrens auftritt (Art. 12–15 DSGVO).[726] Darüber hinaus muss die Möglichkeit verbleiben, eine **Berichtigung** unrichtiger Daten vorzunehmen (Art. 16 DSGVO).[727]

b. Handlungsoptionen und -empfehlungen

Zur Einhaltung der datenschutzrechtlichen Vorgaben sind vielfältige Optionen denkbar, die unter anderem auch mit der Art und Weise der Ausgestaltung der

[721] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 102f., 330.
[722] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 98ff., 252ff., 265ff.
[723] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 96ff., 252ff., 265ff., 325.
[724] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 93, 98ff., 252ff., 311ff., 322ff.
[725] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 98ff., 252ff., 311ff., 322ff.
[726] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 99ff., 252ff., 265ff., 310f., 321ff.
[727] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 252ff., 265ff., 311, 321ff.

Entwicklung und des Betriebs des informationstechnischen Systems in engem Zusammenhang stehen. Sofern nicht der „Königsweg“ des Betriebs durch die Justiz selbst beschritten werden sollte, sondern eine Kooperation mit privaten Dritten erfolgt, ist die Einhaltung des Datenschutzrechts gleichwohl sicherzustellen. Im Rahmen der Kooperation mit Unternehmen für die Bereitstellung der informationstechnischen Systeme sollte darauf geachtet werden, dass die gesetzlichen Vorgaben auch durch den Kooperationspartner beachtet werden, einschließlich Kontroll- und Sanktionsmöglichkeiten.

aa. Rechtsgrundlage der Datenverarbeitung

Für die Datenverarbeitung von informationstechnischen Systemen bei der richterlichen Entscheidungsfindung bedarf es einer entsprechenden rechtlichen Grundlage.[728] Es besteht **prinzipiell eine solche Rechtsgrundlage** für die Verarbeitung personenbezogener Daten in diesem Kontext gem. **§ 3 BDSG**, sowohl für den entscheidungs*ersetzenden*[729] als auch für den entscheidungs*unterstützenden*[730] Einsatz.[731] Durch diese Norm wird öffentlichen Stellen – Rechtsprechungsorgane des Bundes und der Länder gem. § 2 Abs. 1 und Abs. 2 BDSG eingeschlossen – im Rahmen ihrer Befugnisse die Datenverarbeitung, die zur Erfüllung ihrer Aufgaben erforderlich ist, ermöglicht.[732] Hinsichtlich des entscheidungsersetzenden Einsatzes informationstechnischer Systeme müssten zur Erfüllung der Voraussetzungen des § 3 BDSG erst entsprechende Aufgabenzuordnungs- bzw. Zuständigkeitsregelungen im Prozessrecht (siehe hierzu vertieft Abschnitt C.II.6.a) geschaffen werden.[733] Für den entscheidungsunterstützenden Bereich ist dies indes nicht notwendig, da diese Zuständigkeitsbestimmungen bereits existieren.[734]
Neben der Schaffung einer Aufgabenzuordnungs- bzw. Zuständigkeitsregelung im Prozessrecht ist aufgrund des grundsätzlichen Verbots automatisierter Entscheidungsfindung nach Art. 22 Abs. 1 DSGVO für den entscheidungse*rsetzenden* Einsatz informationstechnischer Systeme eine **spezielle rechtliche Grundlage für die automatisierte Datenverarbeitung** erforderlich.[735]
Eine **Handlungsoption** eröffnet hierbei Art. 22 Abs. 2 lit. b) DSGVO, der die Schaffung einer speziellen unionalen oder mitgliedstaatlichen Rechtsvorschrift vorsieht, die den automatisierten Entscheidungsvorgang explizit gestattet.[736] Da die Bestimmung besondere datenschützende Anforderungen an diese Gestattung stellt, genügt § 3 BDSG insoweit nicht. Will man daher einen entscheidungs*ersetzenden* Einsatz informationstechnischer Systeme rechtmäßig über diese Option gestalten,

[728] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 95f., 97ff., 101f., 252ff., 265ff., 307ff., 315ff., 320, 323ff., 331.
[729] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 97, 253f., 265ff., 320f., 325.
[730] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 101, 253f., 265ff., 331.
[731] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 97, 101, 253f., 265ff., 320f., 325, 331.
[732] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 97, 253f., 265ff., 320f., 331; Simitis/Hornung/Spiecker gen. Döhmann/*Roßnagel*, 2. Aufl. 2025, BDSG § 3 Rn. 11.
[733] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 97, 253f., 265ff., 325.
[734] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 101.
[735] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 97, 315ff., 323ff., 331.
[736] Die KI-VO bildet keine solche Rechtsgrundlage, ErwGr. 140 S. 3 KI-VO.

müsste der Gesetzgeber diese automatisierte Entscheidungsfindung in einer Rechtsgrundlage unter Einhaltung der besonderen Anforderungen des Art. 22 Abs. 2 lit. b DSGVO zunächst gestatten. Eine solche Rechtsgrundlage müsste angemessene Maßnahmen zur Wahrung der Rechte und Freiheiten sowie der berechtigten Interessen der betroffenen Person enthalten.
Als solche **angemessenen Maßnahmen** können die Maßnahmen in Art. 22 Abs. 3 DSGVO angesehen werden.[737] Namentlich sind dies insbesondere das Recht auf Erwirkung des Eingreifens einer Person seitens des Verantwortlichen, auf Darlegung des eigenen Standpunkts und auf Anfechtung der Entscheidung (worunter auch eine nächste Instanz vor einem menschlichen Richter fällt[738]; zur Thematik der Verfügbarkeit von Rechtsbehelfen siehe Abschnitt C.II.7), sowie im Rückgriff auf ErwGr 71 S. 4 der DSGVO die spezifische Unterrichtung der betroffenen Person sowie die Erläuterung der nach einer entsprechenden Bewertung getroffenen Entscheidung, wobei festzuhalten ist, dass Art. 22 Abs. 2 lit. b) DSGVO keine abschließenden Vorgaben ausweist.[739] Eine weitere mögliche angemessene Maßnahme im Sinne des Art. 22 Abs. 2 lit. b) DSGVO ist die Offenlegung des Programmcodes, obgleich diese Maßnahme regelmäßig in ihrer Intensität jenseits des vom Unionsgesetzgeber intendierten Ziels anzusiedeln sein dürfte und mit Sicherheitsinteressen konfligiert (siehe hierzu bereits im Abschnitt C.II.4.b.cc(1)).[740] Eine an diesem gemessen maßvollere Maßnahme dürfte die Gewährung der Einsicht in die Bewertungsmaßstäbe bzw. der Erhalt aussagekräftiger Informationen über die involvierte Logik des Systems sein.[741] Vor dem Hintergrund des ErwGr. 71 UAbs. 2 der DSGVO können als angemessene Maßnahmen zudem das Verwenden geeigneter mathematischer oder statistischer Verfahren sowie das Treffen technischer und organisatorischer Maßnahmen zur Fehlerrisikominimierung, Fehlerkorrektur sowie Diskriminierungsverhinderung erwogen werden.[742] Beim Ergreifen der Maßnahmen sollte insgesamt beachtet werden, dass diese den Mitgliedstaat und nicht, wie die Maßnahmen gem. Art. 22 Abs. 3 i. V. m. Abs. 2 lit. a und lit. c DSGVO, den Verantwortlichen adressieren und bei der Ausgestaltung der Maßnahmen dem Mitgliedstaat in der Folge ein größerer Spielraum zukommt.[743] Die Nutzung der Öffnungsklausel des Art. 22 Abs. 2 lit. b) DSGVO ist dem deutschen Gesetzgeber auch nicht unbekannt, es existieren bereits mehrere entsprechende nationale Normen.[744] Auf die im Rahmen ihrer Implementierungen gesammelten Erfahrungen sollte im Zuge der Ausgestaltung der Rechtsgrundlage für den entscheidungsersetzenden Einsatz informationstechnischer Systeme daher auch Rückgriff genommen werden. Darüber hinaus überschneiden sich die Zielrichtungen, die Art. 22 DSGVO setzt, mit Zielen, die sich auch bereits aus vorstehend betrachteten rechtlichen Verbürgungen ergeben, so dass sich die Anforderungen bei Schaffung einer gesetzlichen Grundlage problemlos abbilden ließen.
Eine weitere **Handlungsoption** zur Ermöglichung automatisierter Entscheidungsfindung stellt der **Rekurs auf die Einwilligung** beider Parteien in die

[737] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 317f. m. w. N.; vgl. auch Taeger/Gabel/*Taeger*, 5. Aufl. 2026, DS-GVO Art. 22 Rn. 66.
[738] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 323f.
[739] BeckOK DatenschutzR/*von Lewinski*, 57. Ed. 1.8.2026, DS-GVO Art. 22 Rn. 47.
[740] Paal/Pauly/*Martini*, 4. Aufl. 2026, VO (EU) 2016/679 Art. 22 Rn. 36.
[741] Vgl. Paal/Pauly/*Martini*, 4. Aufl. 2026, VO (EU) 2016/679 Art. 22 Rn. 36a m. w. N.
[742] Simitis/Hornung/Spiecker gen. Döhmann/*Scholz*, 2. Aufl. 2025, DS-GVO Art. 22 Rn. 50.
[743] Paal/Pauly/*Martini*, 4. Aufl. 2026, VO (EU) 2016/679 Art. 22 Rn. 35a.
[744] Vgl. nur Simitis/Hornung/Spiecker gen. Döhmann/*Scholz*, 2. Aufl. 2025, DS-GVO Art. 22 Rn. 51ff. m. w. H.

automatische Datenverarbeitung gem. Art. 22 Abs. 2 lit. c) DSGVO dar. Diese Option kann alternativ oder zusätzlich zum Erlass einer Rechtsgrundlage im Sinne von Art. 22 Abs. 2 lit. b) DSGVO verfolgt werden und beispielsweise auf bestimmte Bereiche beschränkt oder übergangsweise ermöglicht werden, etwa zur – aufgrund der (freiwilligen) Einwilligung – mit Blick auf Akzeptanzthematiken niedrigschwelligen Erprobung von entscheidungsersetzenden informationstechnischen Systemen. Freilich sind im Fall des Art. 22 Abs. 2 lit. c) DSGVO (ebenfalls – und hier sogar zwingend –) die erforderlichen Maßgaben gem. Art. 22 Abs. 3 DSGVO zu beachten.
Darüber hinaus lässt Art. 23 DSGVO die Beschränkung von Rechten (und Pflichten) der DSGVO zu. Dieser umfasst ausweislich seines Wortlauts insbesondere auch Regelungen des Art. 22 DSGVO. Art. 23 DSGVO erfordert eine unionale oder mitgliedstaatliche Rechtsvorschrift, die unter Wahrung der Wesensgehalte der Grundrechte sowie Grundfreiheiten und der Verhältnismäßigkeit zumindest eines der in Art. 23 Abs. 1 DSGVO abschließend aufgeführten Rechtsgüter sicherstellt; im vorliegenden Kontext relevant sind insbesondere Art. 23 Abs. 1 lit. f (Schutz der Unabhängigkeit der Justiz und Schutz von Gerichtsverfahren) und lit. j (Durchsetzung zivilrechtlicher Ansprüche) DSGVO.[745] Die beschränkenden Normen müssen die Mindestanforderungen aus Art. 23 Abs. 2 DSGVO kumulativ wahren.[746]
Art. 23 DSGVO ermöglicht weitere Beschränkungen der von ihm erfassten Rechte (und Pflichten), über die in den jeweiligen Artikeln der DSGVO erfassten Ausnahmen und Beschränkungen hinaus, wobei bei einer Überschneidung von deren Zwecken mit Art. 23 DSGVO die Möglichkeiten des Art. 23 DSGVO insoweit regelmäßig gesperrt sein dürften.[747] Infolge von Beschränkungen darf das Regel-Ausnahme-Verhältnis zwischen Recht (= Regel) und Beschränkungen (= Ausnahme) nicht umgekehrt werden, zudem sind die beschränkenden Regelungen restriktiv auszulegen.[748] Als **Handlungsoption** kann die weitestgehende Restriktion der Betroffenenrechte im Rahmen der Ermöglichung des entscheidungsersetzenden informationstechnischen Systems durch Überwindung des Verbots der automatisierten Entscheidung gem. Art. 22 Abs. 1 DSGVO mittels der Ausnahmeregelungen nach Art. 22 Abs. 2 DSGVO, insbesondere Art. 22 Abs. 2 lit. b) DSGVO ergriffen werden. Dies könnte flankiert werden durch eine Restriktion auch der weiteren Betroffenenrechte nach Art. 12 ff. DSGVO mittels einer beschränkenden Regelung nach Art. 23 DSGVO (zu den Handlungsoptionen bezüglich der Betroffenenrechte im datenschutzrechtlichen Kontext siehe noch nachfolgend). Die Beschränkungsmöglichkeit ist indes fakultativer Natur, es besteht dem Grunde nach – abgesehen von den Ausnahmefällen von Beschränkungsgeboten aufgrund primären Unionsrechts oder mitgliedstaatlichen Verfassungsrechts – keine Pflicht, eine solche Beschränkungsregelung zu erlassen.[749] Insoweit besteht als weitere **Handlungsoption**, keine die Betroffenenrechte weiter zurückdrängenden Regelungen im Kontext des Art. 22 DSGVO zur Ermöglichung des Einsatzes entscheidungsersetzender informationstechnischer Systeme (sowie bezüglich anderer Betroffenenrechte) zu erlassen, wofür insbesondere Vertrauensgesichtspunkte angeführt werden können (zu möglichen Vorteilen unterschiedlich intensiven Maßes an Datenschutz siehe nachfolgend). Freilich können und sollten die Vor- und Nachteile des Zulassens und Beschränkens der unterschiedlichen Datenschutzrechte in

[745] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 318, 323ff.
[746] BeckOK DatenschutzR/*Stender-Vorwachs/Wolff/von Ungern-Sternberg*, 57. Ed. 1.8.2026, DS-GVO Art. 23 Rn. vor 1.
[747] Simitis/Hornung/Spiecker gen. Döhmann/*Dix*, 2. Aufl. 2025, DS-GVO Art. 23 Rn. 1f.
[748] Simitis/Hornung/Spiecker gen. Döhmann/*Dix*, 2. Aufl. 2025, DS-GVO Art. 23 Rn. 1ff.
[749] Kühling/Buchner/*Bäcker*, 4. Aufl. 2024, DS-GVO Art. 23 Rn. 3.

verschiedener Kombination sowie Intensität betrachtet sowie evaluiert und der Gesamtkomplex in einen bestmöglichen Ausgleich gebracht bzw. einem bestmöglichen Resultat zugeführt werden.
Mit Blick auf die vorstehenden Handlungsoptionen ist **zu empfehlen**, einen entscheidungs*ersetzenden* Einsatz auf eine gesetzliche Grundlage zu stellen und Art. 22 Abs. 2 lit. b DSGVO heranzuziehen. Eine gesetzliche Grundlage muss für den Einsatz solcher Systeme ohnehin geschaffen werden, so dass die datenschutzrechtlichen Aspekte gleich mitgeregelt werden können. Darüber hinaus hat diese Handlungsoption den Vorteil, dass nicht stets das Einverständnis aller beteiligter Parteien für die Nutzung der Systeme erforderlich ist, so dass die Effizienzgewinne nicht durch taktische Verweigerungen der Einwilligung torpediert werden können.

bb. Gewährleistung der Datensicherheit

Die Gewährleistung von Datensicherheit bei der Verarbeitung personenbezogener Daten ist ein essentieller Aspekt des Datenschutzrechts. In diesem Zusammenhang sind insbesondere der DSGVO Anforderungen an den Verantwortlichen für die Verarbeitung personenbezogener Daten zu entnehmen.
Dieser Pflichtenkreis umfasst gem. Art. 24 DSGVO unter Berücksichtigung eines risikobasierten Ansatzes geeignete *technische und organisatorische Maßnahmen*, um die DSGVO-Anforderungen einzuhalten und dies sicherzustellen.[750] Derartige Maßnahmen sind alle erforderlichen Maßnahmen, um die Beachtung von Datenschutz und -sicherheit bei der Erhebung, Verarbeitung und Nutzung personenbezogener Daten sowie den dazu betriebenen Verfahren sicherzustellen.[751] Als Maßnahmentypen können hierfür frühzeitige Pseudonymisierungen, Maßnahmen der Datenminimierung sowie Maßnahmen zur Ermöglichung der Überwachung der Datenverarbeitung durch den Einzelnen und die Implementierung sowie Verbesserung von Sicherheitsfunktionen herangezogen werden.[752] Als technische Maßnahmen können u. a. implementiert werden Zugriffskontrollen, Weitergabekontrollen oder Verschlüsselungen, aber auch bauliche Maßnahmen, die unbefugtem Zugriff vorbeugen, rein physisch wirkende Vorkehrungen, wie das Wegschließen von Datenträgern, aber auch Inventuren sowie Klassifikationen von Datenbeständen und Analysen von Verarbeitungsverfahren.[753] Organisatorische Maßnahmen zur Sicherstellung der Datensicherheit sind Protokollierung, Schulungen der Mitarbeiter oder das Vieraugenprinzip, sowie Stichprobenroutinen oder Verpflichtungserklärungen der Mitarbeiter vergleichbar § 5 S. 2 BDSG a. F., der Aufbau einer Datenschutzorganisation oder die Bestellung eines Datenschutzbeauftragten (letzterer ist aufgrund der Ausnahme in Art. 37 Abs. 1 lit. a DSGVO nicht verpflichtend für Gerichte, die innerhalb ihrer justiziellen Tätigkeit handeln).[754] Die Abgrenzung der

[750] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 318f.; BeckOK DatenschutzR/*Schmidt/Brink*, 57. Ed. 1.8.2025, DS-GVO Art. 24 vor Rn. 1.
[751] BeckOK DatenschutzR/*Schmidt/Brink*, 57. Ed. 1.8.2025, DS-GVO Art. 24 Rn. 14.
[752] Paal/Pauly/*Martini*, 4. Aufl. 2026, VO (EU) 2016/679 Art. 24 Rn. 23 mit Hinweis auf Erwägungsgrund 78 S. 3 der DSGVO.
[753] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 318f.; BeckOK DatenschutzR/*Schmidt/Brink*, 57. Ed. 1.8.2025, DS-GVO Art. 24 Rn. 15; Paal/Pauly/*Martini*, 4. Aufl. 2026, VO (EU) 2016/679 Art. 24 Rn. 21.
[754] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 318f.; BeckOK DatenschutzR/*Schmidt/Brink*, 57. Ed. 1.8.2025, DS-GVO Art. 24 Rn. 15; Taeger/Gabel/*Lang*, 5. Aufl. 2026, DS-GVO Art. 24 Rn. 25; Paal/Pauly/*Martini*, 4. Aufl. 2026, VO (EU) 2016/679 Art. 24 Rn. 21f.; Simitis/Hornung/Spiecker gen. Döhmann/*Drewes*, 2. Aufl. 2025, DS-GVO Art. 37 Rn. 10.

Maßnahmen zwischen organisatorischer oder technischer Natur ist weder trennscharf möglich noch nötig; vielmehr ist die Eignung dazu entscheidend, die Verarbeitung nach den gesetzlichen Vorgaben des Datenschutzes sicherzustellen und einen entsprechenden Nachweis zu ermöglichen sowie die Tatsache, dass die Ausgestaltung der Vielzahl von Maßnahmen im Einzelfall variiert und es **keinen verbindlichen Mindeststandard** gibt.[755]
Eine Orientierung an Katalogen mit **Referenzmaßnahmen** ist hingegen möglich und durchaus sinnvoll und sollte daher auch erfolgen; ein solcher ist u. a. das **Standard-Datenschutzmodell der Datenschutzbehörden von Bund und Ländern**.[756] Dieses kann aufgrund seiner Aktualisierungen und der Ausgestaltung, dass es die rechtlichen Anforderungen der DSGVO mit Hilfe von Gewährleistungszielen in technische sowie organisatorische Maßnahmen übersetzt, von Vorteil sein.[757] Auch wird die konkrete Praxisanwendung durch das Zurverfügungstellen veröffentlichter Regelungen, sogenannter „Bausteine“, unterstützt.[758] Eine Konkretisierung der Pflichten aus Art. 24 DSGVO findet sich insbesondere in Art. 25, 32 und 35 DSGVO.[759] Hinsichtlich der Datensicherheit sind hierbei vor dem Hintergrund des Art. 5 Abs. 1 lit. f DSGVO, der die Integrität und Vertraulichkeit der Datenverarbeitung als DSGVO-Grundsatz normiert, insbesondere Art. 25 und 32 DSGVO von Bedeutung und näher zu betrachten, die diesen durch die erforderlichen technischen und organisatorischen Maßnahmen konkretisieren.[760]
Art. 25 DSGVO legt bestimmte rechtliche Designpflichten nieder.[761] Anforderungen des Datenschutzes sind nach dem **„privacy by design“-Prinzip** nach Art. 25 Abs. 1 DSGVO bereits im Stadium der Konzeption sowie Entwicklung zu berücksichtigen.[762] Auch besteht die Verpflichtung des Verantwortlichen nach Art. 25 Abs. 1 DSGVO, geeignete technische und organisatorische Maßnahmen zu implementieren, die sicherstellen, dass nur die für den Verarbeitungszweck notwendigen Daten verarbeitet werden, sog. „data protection by default“.[763] Die Maßnahmen sind bereits zum Zeitpunkt der Mittelfestlegung zu treffen, was mit Blick auf Eigenentwicklungen eines Verantwortlichen das **Einfließen konkreter technischer sowie organisatorischer Maßnahmen bereits in der Konzeptphase** bedeutet bzw. bei einer Entwicklung durch einen Dritten die Beachtung der Vorgaben

[755] Taeger/Gabel/*Lang*, 5. Aufl. 2026, DS-GVO Art. 24 Rn. 25ff.; *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 318f.
[756] Taeger/Gabel/*Lang*, 5. Aufl. 2026, DS-GVO Art. 24 Rn. 27.
[757] *Der Bundesbeauftragte für den Datenschutz und die Informationsfreiheit*, Das Standard-Datenschutzmodell, abrufbar unter https://www.bfdi.bund.de/DE/Fachthemen/Inhalte/Technik/SDM.html (Abruf am 08.04.2026, 17:28 Uhr); *Der Landesbeauftragte für Datenschutz und Informationsfreiheit Mecklenburg-Vorpommern*, Das Standard-Datenschutzmodell, abrufbar unter https://www.datenschutz-mv.de/datenschutz/datenschutzmodell/ (Abruf am 08.04.2026, 17:30 Uhr).
[758] Die „Bausteine“ sind abrufbar unter *Der Landesbeauftragte für Datenschutz und Informationsfreiheit Mecklenburg-Vorpommern*, Das Standard-Datenschutzmodell, https://www.datenschutz-mv.de/datenschutz/datenschutzmodell/ (Abruf am 08.04.2026, 17:30 Uhr).
[759] BeckOK DatenschutzR/*Schmidt/Brink*, 57. Ed. 1.8.2025, DS-GVO Art. 24 vor Rn. 1; Paal/Pauly/*Martini*, 4. Aufl. 2026, VO (EU) 2016/679 Art. 25 Rn. 1.
[760] Gola/Heckmann/Rost/*Pötters*, 4. Aufl. 2026, DS-GVO Art. 5 Rn. 30.
[761] Paal/Pauly/*Martini*, 4. Aufl. 2026, VO (EU) 2016/679 Art. 25 vor Rn. 1.
[762] Ehmann/Selmayr/*Baumgartner*, 3. Aufl. 2024, DS-GVO Art. 25 Rn. 1; *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 318f.
[763] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 318f. m. w. N.

des Art. 25 Abs. 1 DSGVO bei dessen Auswahl (insbesondere öffentlichen Vergaben) oder der Erstellung eines Pflichtenkatalogs.[764]
Herausgehoben seien für den richterlichen Bereich die **organisatorischen Maßnahmen** insbesondere regelmäßige Schulungen sowie interne Durchsetzungen – beispielsweise in Form von Ahndung von Datenschutzverstößen –, ebenso wie die Bildung interdisziplinärer Projektteams.[765] Durch die Einbindung von **Experten der verschiedenen, notwendigen Disziplinen** in ein Team kann **bereits ab dem Zeitpunkt der Entwicklung** gesichert werden, dass alle notwendigen Anforderungen sowie Gestaltungsnotwendigkeiten und -wünsche beachtet und im System implementiert werden. Ein derartiges Vorgehen ist auch bereits im Zusammenhang mit der Entwicklung von informationstechnischen Systemen für die richterliche Entscheidungsfindung bekannt und bewährt – neben der Entwicklung der dem vorliegenden Forschungsprojekt entspringenden MAKI-KI wurde ein solches Vorgehen beispielsweise auch im Rahmen der Entwicklung des Oberlandesgerichtsassistenten „OLGA“, der die Richter in den Diesel-Senaten des OLG Stuttgart unterstützt, als vorteilhaft identifiziert.[766] Auch die **stringente Durchsetzung von Maßnahmen bei Datenschutzverstößen** kann die Beachtung der Datenschutzanforderungen fördern und mit regelmäßigen Schulungen die Kenntnis und den richtigen Umgang mit den Systemen verfestigen.
Gerade auf die Beachtung und Erfüllung der Anforderungen des Art. 25 DSGVO sollte Wert gelegt werden. Aufsichtsbehörden haben bereits angekündigt, die Einhaltung der Anforderungen des Art. 25 DSGVO stärker zu überprüfen, wobei jedoch zu beachten ist, dass die Aufsichtsbehörden nicht für die Aufsicht über die Verarbeitungen zuständig sind, die Gerichte im Rahmen ihrer justitiellen Tätigkeit vornehmen, mithin Maßnahmen, die in richterlicher Unabhängigkeit erbracht werden, Art. 55 Abs. 3 DSGVO.[767] Dies umfasst insbesondere den Entscheidungsspruch einschließlich Sach- und Verfahrensentscheidungen, die diesem dienen oder diesen vor- bzw. nachbereiten; nicht umfasst sind hingegen u. a. Justizverwaltungsangelegenheiten.[768] Damit sollte keine Kontrollfreiheit vorgesehen werden, sondern vielmehr eine Selbstkontrolle der Justiz in diesem Bereich erfolgen, ErwGr. 20 der DSGVO.[769] Diese Selbstkontrolle hat jedoch weder im Verordnungstext selber Niederschlag gefunden, noch in der praktischen Umsetzung auf Bundes- oder Länderebene.[770] Jedenfalls aus praktischer Sicht erscheint aufgrund bereits erfolgender Beschwerden gegen gerichtliche Datenverarbeitungen, dass eine solche Implementierung sinnvoll ist; mit Blick auf den KI-Einsatz wird die Thematik präsumtiv umso relevanter.[771] Daher sollte als Handlungsoption in Betracht gezogen werden,

[764] BeckOK DatenschutzR/*Paulus/Schmidt*, 57. Ed. 1.8.2025, DS-GVO Art. 25 Rn. 4a ff.; Ehmann/Selmayr/*Baumgartner*, 3. Aufl. 2024, DS-GVO Art. 25 Rn. 17.
[765] Gola/Heckmann/Rost/*Werkmeister*, 4. Aufl. 2026, DS-GVO Art. 25 Rn. 17.
[766] *Spoenle*, OLGA – der KI-Assistent für Dieselverfahren, in: Reuß/Laß (Hrsg.), Göttinger Kolloquien zur Digitalisierung des Zivilverfahrensrechts Kolloquien im Sommersemester 2022 | Kolloquien im Wintersemester 2022/2023, Göttingen 2024, S. 175.
[767] Simitis/Hornung/Spiecker gen. Döhmann/*Polenz*, 2. Aufl. 2025, DS-GVO Art. 55 Rn. 19f.; Ehmann/Selmayr/*Baumgartner*, 3. Aufl. 2024, DS-GVO Art. 25 Rn. 6 m. w. N.; es wird angenommen, dass die ohnehin schon praktisch hoch bedeutsame Norm Art. 25 DSGVO i. H. a. KI-Systeme einen erheblichen Bedeutungszuwachs erfährt und zu „der“ zentralen Norm der DSGVO im Rahmen der KI-Regulierung aufsteigt, Ehmann/Selmayr/*Baumgartner*, 3. Aufl. 2024, DS-GVO Art. 25 Rn. 5f.
[768] Simitis/Hornung/Spiecker gen. Döhmann/*Polenz*, 2. Aufl. 2025, DS-GVO Art. 55 Rn. 20.
[769] Paal/Pauly/*Körffer*, 4. Aufl. 2026, VO (EU) 2016/679 Art. 55 Rn. 6.
[770] Paal/Pauly/*Körffer*, 4. Aufl. 2026, VO (EU) 2016/679 Art. 55 Rn. 6.
[771] Unabhängig von der bisweilen bejahten Frage, ob eine Verpflichtung zur Umsetzung aus dem Europäischen Primärrecht folgt. Ehmann/Selmayr/*Selmayr*, 3. Aufl. 2024, DS-GVO Art. 55 Rn. 15; Paal/Pauly/*Körffer*, 4. Aufl. 2026, VO (EU) 2016/679 Art. 55 Rn. 6.

eine entsprechende Aufsichtsstruktur aufzubauen. Der hierbei gegebene Spielraum zur Einrichtung könnte u. a. durch einen von Richtern gewählten, unabhängigen Kontrollmechanismus[772] oder durch ein spezielles Datenschutz-Aufsichts-Gericht, ausgefüllt werden.[773]
Art. 32 DSGVO konkretisiert Art. 24 DSGVO sowie Art. 5 Abs. 1 lit. f DSGVO. Er verpflichtet umfassend den Datenverarbeitungsverantwortlichen sowie den Auftragsverarbeiter, geeignete Maßnahmen zu ergreifen, um die **Sicherheit der Datenverarbeitung** zu gewährleisten und gibt den Verpflichteten das regulatorische Rahmenwerk zur Umsetzung derartiger Maßnahmen.[774] Die Regelungen der DSGVO bleiben von der KI-VO unberührt und werden durch diese mit Blick auf KI-Systeme ergänzt; die Anforderungen nach Art. 32 DSGVO gelten auch für KI-Systeme unverändert.[775] Einem sog. relativen Ansatz folgend, sind bei der Auswahl der Maßnahmen Schutzmaßnahmen und Risiko einander gegenüberzustellen und eine Balance von Schutzniveau, Stand der Technik sowie Risiko zu finden.[776] Im Rahmen der Umsetzungspflicht ist eine Verhältnismäßigkeitsprüfung vorzunehmen, für die Art. 32 Abs. 1 DSGVO die zu berücksichtigenden Elemente vorgibt.[777]
Der Stand der Technik kann hierbei verstanden werden als bekannte, bewährte, effektive und nicht veraltete, auf dem Markt verfügbare Maßnahmen.[778] Die aktuelle Marktsituation und Empfehlungen sowie Informationen von staatlichen Stellen, wie etwa Veröffentlichungen zum IT-Grundschutz des Bundesamtes für Sicherheit in der Informationstechnik (BSI), aber auch die enge Einbindung des BSI in das gesamte Verfahren sind wichtige Faktoren, die hierbei Beachtung und als Handlungsoptionen in die Umsetzung Einzug finden sollten.[779] Bei den Implementierungskosten ist der finanzielle Aufwand für die Umsetzung zu beachten und, dass kein unbegrenzter Aufwand vonnöten ist.[780] Die Brücke zum risikobasierten Ansatz des Art. 24 Abs. 1 DSGVO schlagend, sind auch zu berücksichtigen: Art, Umfang, Umstände sowie Zwecke der Verarbeitung und Schwere sowie die Eintrittswahrscheinlichkeit für die Freiheiten und Rechte natürlicher Personen im Einzelfall.[781] Mögliche Maßnahmen und Schutzziele weist Art. 32 Abs. 1 Hs. 2 DSGVO nicht abschließend aus; regelmäßig werden zusätzliche Maßnahmen notwendig sein.[782] Für die Beurteilung des angemessenen Schutzniveaus listet Art. 32 Abs. 2 DSGVO einen nicht abschließenden Katalog an Risiken auf.[783] Art. 32 Abs. 4 DSGVO verpflichtet Verantwortliche sowie Auftragsverarbeiter grundsätzlich dazu, dafür zu sorgen, dass ihnen jeweils unterstelltes Personal mit Zugang zu personenbezogenen Daten bei deren

[772] Etwa nach dem Vorbild eines derartigen Mechanismus beim EuGH (Unabhängiger Dreierausschuss, dessen Mitglieder aus dem Kreis der Richter und Generalanwälte ausgewählt und auf Vorschlag des Präsidenten des Gerichtshofs für die Dauer von dessen Amtszeit vom Gerichtshof ernannt werden), Ehmann/Selmayr/*Selmayr*, 3. Aufl. 2024, DS-GVO Art. 55 Rn. 15 m. w. H.
[773] Ehmann/Selmayr/*Selmayr*, 3. Aufl. 2024, DS-GVO Art. 55 Rn. 15 m. w. H.
[774] Ehmann/Selmayr/*Hladjk*, 3. Aufl. 2024, DS-GVO Art. 32 Rn. 1, 3.
[775] Paal/Pauly/*Martini*, 4. Aufl. 2026, VO (EU) 2016/679 Art. 32 Rn. 17k.
[776] BeckOK DatenschutzR/*Paulus*, 57. Ed. 1.11.2021, DS-GVO Art. 32 Rn. 7.
[777] Gola/Heckmann/Rost/*Piltz*, 4. Aufl. 2026, DS-GVO Art. 32 Rn. 9, 12ff.
[778] Taeger/Gabel/*Bock*, 5. Aufl. 2026, DS-GVO Art. 32 Rn. 25; Gola/Heckmann/Rost/*Piltz*, 4. Aufl. 2026, DS-GVO Art. 32 Rn. 18.
[779] Gola/Heckmann/Rost/*Piltz*, 4. Aufl. 2026, DS-GVO Art. 32 Rn. 19.
[780] Gola/Heckmann/Rost/*Piltz*, 4. Aufl. 2026, DS-GVO Art. 32 Rn. 20f.
[781] Gola/Heckmann/Rost/*Piltz*, 4. Aufl. 2026, DS-GVO Art. 32 Rn. 22.
[782] Ehmann/Selmayr/*Hladjk*, 3. Aufl. 2024, DS-GVO Art. 32 Rn. 28.
[783] Taeger/Gabel/*Bock*, 5. Aufl. 2026, DS-GVO Art. 32 Rn. 43 m. w. H.

Verarbeitung nur gemäß den Anweisungen des Verantwortlichen handelt.[784] Hiermit sollen Betroffene vor Grundrechtsverletzungen geschützt werden; Schutzmaßnahmen vor zweckfremder Verarbeitung sowie Rechte- und Rollenkonzepte auf der Basis der Erforderlichkeit kommen hierbei u. a. in Betracht.[785] Für den Einsatz informationstechnischer Systeme bei der richterlichen Entscheidungsfindung sind verschiedene Ausgestaltungsmöglichkeiten denkbar – so sind beispielsweise die Systeme entscheidungsersetzend sowie entscheidungsunterstützend und jeweils in Form eines regelbasierten Systems, Systems maschinellen Lernens bzw. als Mischsystem denkbar. Zudem sind insbesondere dem entscheidungsunterstützenden Einsatz informationstechnischer Systeme viele verschiedene Funktionalitäten zuzuordnen.[786] Hinzu kommen weitere Aspekte wie etwa die Frage der Ansiedelung der informationstechnischen Systeme (siehe hierzu Abschnitt C.II.2.a.bb(2)). Die jeweiligen zu adressierenden Risiken und dafür notwendigen Maßnahmen sind daher abgestimmt auf den konkreten jeweiligen Fall zu bestimmen, wobei beachtet werden sollte, dass gerade im justiziellen Kontext intime sowie schutzbedürftige, sensible Informationen verarbeitet werden.[787] Die notwendigen Maßnahmen sind mithin mit Blick auf die konkrete Ausgestaltung des informationstechnischen Systems, dessen Ansiedelung sowie dessen Funktionalität zu bestimmen. Bei der Umsetzung der Vorgaben des Art. 32 DSGVO sollte hierbei das Synergiepotential genutzt werden, das durch den engen Zusammenhang u. a. mit Art. 15 KI-VO sowie Art. 9 KI-VO (siehe hierzu bereits zuvor im Abschnitt C.II.3.b.bb) besteht.

Art. 35 DSGVO regelt die **Datenschutz-Folgenabschätzung (DSFA)**, mithin ein Verfahren, mittels dessen eine bestimmte Datenverarbeitung, die vermutlich zu einem hohen Risiko für Betroffene führt, bezeichnet, ihre Erforderlichkeit und Verhältnismäßigkeit bewertet und letztlich ihre Risiken für die Betroffenen durch die Identifizierung von Gegenmaßnahmen zumindest abgemildert wird.[788] Nach diesem Konzept hat der Verantwortliche eine detaillierte Risikoanalyse in eigener Verantwortung durchzuführen und bei Vorliegen eines hohen Risikos eine DSFA – vor Beginn der jeweiligen Verarbeitungstätigkeit und bestenfalls bereits in der Entwicklungsphase – durchzuführen; die zukünftige Praxisrelevanz der Norm mit Blick auf den Einsatz informationstechnischer Systeme erscheint naheliegend.[789] Das Synergiepotential aufgrund der großen Überschneidungen zwischen dem Risikomanagement gem. KI-VO sowie der DSGVO ist augenfällig und sollte daher genutzt werden; die DSFA im Kontext von KI bilden eine effiziente Möglichkeit zur Umsetzung dieser KI-VO-Vorgaben.[790] Auf die DSFA nimmt die KI-VO auch an mehreren Stellen, z. B. Art. 26 Abs. 9 KI-VO und Art. 27 Abs. 4 KI-VO Bezug und im Rahmen der EU-Konformitätserklärung nach Art. 47 KI-VO i. V. m. Anhang V Nr. 5 zur KI-VO muss die Erklärung der DSGVO-Konformität enthalten sein.[791]

---

[784] Ehmann/Selmayr/*Hladjk*, 3. Aufl. 2024, DS-GVO Art. 32 Rn. 16; Taeger/Gabel/*Bock*, 5. Aufl. 2026, DS-GVO Art. 32 Rn. 45.
[785] Taeger/Gabel/*Bock*, 5. Aufl. 2026, DS-GVO Art. 32 Rn. 45 m. w. H.
[786] Vgl. nur *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 35ff.
[787] Vgl. nur *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 82, 309f.
[788] Die Norm vervollständigt die Anforderungen an den Verantwortlichen gem. Art. 24, 25, 32 DSGVO, Ehmann/Selmayr/*Baumgartner*, 3. Aufl. 2024, DS-GVO Art. 35 Rn. 1.
[789] Ehmann/Selmayr/*Baumgartner*, 3. Aufl. 2024, DS-GVO Art. 35 Rn. 1f.
[790] Vgl. *Bayerisches Landesamt für Datenschutzaufsicht*, KI & Datenschutz, abrufbar unter https://www.lda.bayern.de/de/ki.html (Abruf am 14.04.2026, 17:05 Uhr).
[791] BeckOK KI-Recht/*Kaufmann*, 7. Ed. 25.7.2026, KI-VO Art. 4 Rn. 47; die Erzeugung länderübergreifender Synergieeffekte befürwortend *E-Justice-Rat*, Strategie für den Einsatz von

Die Absätze innerhalb des Art. 35 DSGVO fügen sich wie folgt ineinander: Hinsichtlich der Frage des „Ob“ einer DSFA können Aufsichtsbehörden sogenannte Positivlisten nach Abs. 4 bzw. Negativlisten nach Abs. 5 über konkrete Verarbeitungsvorgänge erlassen, in denen eine DSFA durchgeführt werden muss bzw. nicht erforderlich sein soll.[792] Fällt der zu betrachtende Vorgang nicht unter (die Listen nach) Abs. 4 und 5 (etwa weil schon keine Aufsichtsbehörde besteht), kann auf die Regelbeispiele zurückgegriffen werden, die Abs. 3 (als lex specialis vorrangig vor Abs. 1 zu prüfen) ausweist und bei deren Einschlägigkeit eine DSFA als erforderlich anzusehen ist.[793] Ist keiner der Tatbestände des Abs. 3 erfüllt, ist Abs. 1 zu prüfen, bei der für die Frage des „Ob“ einer DSFA gem. Art. 35 Abs. 1 S. 1 DSGVO bei der Begutachtung, ob eine Form der Verarbeitung ein hohes Risiko für die Rechte und Freiheiten natürlicher Personen zur Folge hat insbesondere die Topoi der Verwendung neuer Technologien, Art, Umfang sowie Umstände und der Zwecke der Verarbeitung relevant sind. Die Methodik zur Risikofeststellung obliegt dem Verantwortlichen festzulegen; neben einer in der Praxis hochkomplexen Risikoanalyse mittels Matrix kann das Risiko insbesondere unter dem Gesichtspunkt, wie viele der Kriterien aus „WP 248“ der Leitlinien zur Datenschutz-Folgeabschätzung der EU-Aufsichtsbehörden vorliegen, betrachtet werden, wobei hiernach – bei nicht gänzlich starrer Anwendung – das Vorliegen von mindestens zwei der genannten Kriterien für einen Verarbeitungsvorgang regelmäßig ein hohes Risiko anzeigt.[794]
Je nach Kontext können im Rahmen des Einsatzes von informationstechnischen Systemen bei der richterlichen Entscheidungsfindung verschiedene dieser Kriterien in Betracht kommen, als regelmäßig relevant erscheinen insbesondere folgende Kriterien: Automatisierte Entscheidungsfindung; Die Verarbeitung vertraulicher oder höchstpersönlicher Daten, an denen der Betroffene ein besonderes Schutzinteresse hat; Die Verarbeitung im großen Umfang; Die Verarbeitung von Daten Schutzbedürftiger; Die Nutzung eines neuen Prozesses oder einer neuen Technologie – und damit regelmäßig eine Vielzahl der aufgeführten Kriterien.[795]
Vor dem Hintergrund dieses Befundes sollten als Handlungsoptionen eine die gesamte richterliche Sphäre abdeckende Aufsichtsstruktur geschaffen werden sowie die Aufnahme der Verarbeitungen in eine Muss-Liste gem. Art. 35 Abs. 4 DSGVO erfolgen, die im Rahmen des Einsatzes informationstechnischer Systeme in der richterlichen Entscheidungsfindung Relevanz haben.[796] Dies kann der Rechtsklarheit

---

Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 26.

[792] Ehmann/Selmayr/*Baumgartner*, 3. Aufl. 2024, DS-GVO Art. 35 Rn. 4; beispielhaft kann für derartige Positivlisten auf die DSFA Muss-Listen für den nicht-öffentlichen sowie für den öffentlichen Bereich des Landesbeauftragten für den Datenschutz Niedersachsen verwiesen werden, *Der Landesbeauftragte für den Datenschutz Niedersachsen*, Muss-Listen zur Datenschutz-Folgenabschätzung, abrufbar unter https://www.lfd.niedersachsen.de/dsgvo/liste_von_verarbeitungsvorgangen_nach_art_35_abs_4_ds_gvo/muss-listen-zur-datenschutz-folgenabschatzung-179663.html (Abruf am 09.04.2026, 22:25 Uhr).

[793] Simitis/Hornung/Spiecker gen. Döhmann/*Karg*, 2. Aufl. 2025, DS-GVO Art. 35 Rn. 45; Ehmann/Selmayr/*Baumgartner*, 3. Aufl. 2024, DS-GVO Art. 35 Rn. 4.

[794] Taeger/Gabel/*Reibach*, 5. Aufl. 2026, DS-GVO Art. 35 Rn. 16, 18, 20f. jeweils m. w. N.; das „WP 248“ ist abrufbar unter *Artikel-29-Datenschutzgruppe*, WP 248 rev.01, abrufbar unter https://www.datenschutzkonferenz-online.de/media/wp/20171004_wp248_rev01.pdf (Abruf am 14.04.2026, 20:28 Uhr), die Kriterien sind auf Seite 10ff. verzeichnet.

[795] Vgl. Taeger/Gabel/*Reibach*, 5. Aufl. 2026, DS-GVO Art. 35 Rn. 21 m. w. H.

[796] Als zusätzlicher Grund könnten auf den ersten Blick auch die erheblichen Sanktionsmöglichkeiten des Art. 83 Abs. 4 lit. a DSGVO bei Verstößen herangezogen werden. Gem. Art. 83 Abs. 7 DSGVO in Verbindung mit den jeweiligen Bundes- und Ländergesetzen ist eine derartige Haftung insbesondere der Organe der Rechtspflege jedoch ausgeschlossen, siehe hierzu auch noch sogleich; *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter

durch die Klarstellung zuträglich sein, zumal bei Vorliegen der Voraussetzungen gem. Art. 35 Abs. 3 (insb. lit. a und b) respektive 1 DSGVO – das regelmäßig naheliegend scheint – ohnehin eine DSFA erfolgen muss. Die DSFA-Kriterien lässt Art. 35 DSGVO weitestgehend offen, er legt jedoch in Art. 35 Abs. 7 (und 8) DSGVO einen zwingenden Mindestgehalt fest.[797] Als zu beachtende Faktoren werden in veröffentlichten Katalogen in diesem Zusammenhang insbesondere Halluzinationen sowie unrichtige und gebiaste Ausgaben, unzureichende Transparenz, mangelhafte Sicherheitsvorkehrungen, Verlässlichkeit und Robustheit in Bezug auf das KI-Modell, dysfunktionale menschliche Aufsicht und Kontrolle, unzureichende Rechtsgrundlagen und die Sicherstellung von Betroffenenrechten genannt;[798] mithin insbesondere die Aspekte, die als rote Linien des ersten Milestone-Berichts des MAKI-Forschungsprojekts identifiziert wurden.
Die KI-VO sieht – in Kenntnis des Verordnungsgebers von der Figur des Datenschutzbeauftragten – keine Verpflichtung zur Bestellung eines KI-Beauftragten, ähnlich der DSGVO vor; vielmehr sieht die KI-VO eine menschliche Aufsicht nur bei der Verwendung von Hochrisiko-KI-Systemen vor, Art. 14 Abs. 1, 26 Abs. 2 KI-VO.[799] Gleichwohl erscheint gerade mit wachsender Größe einer Institution das Designieren eines entsprechenden Beauftragten bzw. eines beauftragten Gremiums sinnvoll.[800] Die Übertragung der Aufgaben eines „KI-Beauftragten“ auf den Datenschutzbeauftragten erscheint im Hinblick auf Synergiepotential naheliegend; dieser darf andere Pflichten und Aufgaben dem Grunde nach wahrnehmen, Art. 38 Abs. 6 S. 1 DSGVO, soweit keine Interessenskonflikte (bspw. bei Selbstüberwachung und Entscheidung über Mittel sowie Zwecke der Verarbeitung von personenbezogenen Daten) aufkommen und ihm genügend zeitliche Ressourcen für seine Stellung als Datenschutzbeauftragter, vgl. Art. 39 DSGVO, und seine Unabhängigkeit sowie Weisungsfreiheit gewahrt bleibt, vgl. Art. 38 Abs. 3 DSGVO.[801] In größeren Institutionen kann indes die Aufgabenfülle, die im Rahmen des KI-Einsatzes auftreten wird, schnell einen Umfang erreichen, der die zusätzlichen Ressourcen neben der Wahrnehmung der Aufgaben als Datenschutzbeauftragter übersteigt.[802] Auch können mögliche Interessenskonflikte zu Problemen bei einer derartigen Ausgestaltung führen. Daher sollte die Aufgabenübertragung mit Augenmaß behandelt und implementiert werden. Sollten hingegen z. B. durch eine Aufgabenübertragung die zusätzlichen Ressourcen überstiegen werden, ist von einer solchen abzusehen. In einem solchen Falle könnte die Designation eines/einer

https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 26 befürwortet die Erzeugung länderübergreifender Synergieeffekte insbesondere bei der DSFA.

[797] Paal/Pauly/*Martini*, 4. Aufl. 2026, VO (EU) 2016/679 Art. 35 Rn. 44ff.

[798] *Bayerisches Landesamt für Datenschutzaufsicht*, KI & Datenschutz, abrufbar unter https://www.lda.bayern.de/de/ki.html (Abruf am 14.04.2026, 17:05 Uhr).

[799] BeckOK KI-Recht/*Kaufmann*, 7. Ed. 25.7.2026, KI-VO Art. 4 Rn. 42.

[800] Vgl. *Bayerisches Landesamt für Datenschutzaufsicht*, KI & Datenschutz, abrufbar unter https://www.lda.bayern.de/de/ki.html (Abruf am 14.04.2026, 17:05 Uhr).

[801] BeckOK KI-Recht/*Kaufmann*, 7. Ed. 25.7.2026, KI-VO Art. 4 Rn. 48; vgl. auch Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 35f.

[802] Vgl. *Bayerisches Landesamt für Datenschutzaufsicht*, KI & Datenschutz, abrufbar unter https://www.lda.bayern.de/de/ki.html (Abruf am 14.04.2026, 17:05 Uhr).

eigenen KI-Beauftragten verfolgt werden. Diese/r könnte z. B. in die Nähe des Datenschutzbeauftragten angegliedert werden.

cc. Eingabe personenbezogener Daten in Systeme, die nicht der Kontrolle der Justiz unterstehen

Die Eingabe personenbezogener Daten in Systeme, die nicht der Kontrolle der Justiz unterstehen bzw. die datenschutzrechtlichen Anforderungen nicht wahren, z. B. ein cloudbasiertes Dialogsystem wie ChatGPT, ist in der Regel unzulässig.[803] Diesem Befund tritt die Tatsache gegenüber, dass der breite Einsatz von KI im Justizsystem (ohne organisationale Freigabe oder Lizenz) und hierbei insbesondere von ChatGPT bereits gelebte Tatsache ist.[804] Die erhebliche datenschutzrechtliche Problematik dessen ist augenfällig, ebenso wie die datenschutz- und dienstrechtliche Relevanz der Verstöße der im Einzelnen handelnden Person.[805] Die DSGVO enthält selbst verschiedene Möglichkeiten, um auf Verstöße gegen sie zu reagieren. So beinhaltet Art. 82 DSGVO einen weitgehenden, eigenständigen, deliktischen Schadensersatzanspruch mit vermutetem Verschulden für materielle sowie immaterielle Schäden.[806] Zudem stehen den Aufsichtsbehörden neben Untersuchungsbefugnissen nach Art. 58 Abs. 1 DSGVO auch umfangreiche Abhilfebefugnisse gem. Art. 58 Abs. 2 DSGVO zur Verfügung. Flankiert werden diese Befugnisse durch die Möglichkeit von Aufsichtsbehörden gem. Art. 83 DSGVO, empfindliche Geldbußen zu verhängen, die sich grundsätzlich auch gegen Behörden und öffentliche Stellen richten können, Art. 83 Abs. 7 DSGVO e contrario. Gem. Art. 83 Abs. 7 DSGVO besteht indes für im Mitgliedstaat niedergelassene Behörden und öffentliche Stellen des Mitgliedstaates – unabhängig von der Frage der Existenz einer Aufsichtsbehörde im Bereich der justiziellen Gerichtstätigkeit (siehe hierzu bereits zuvor) – eine nationale Öffnungsklausel betreffend das „Ob“ einer Geldbuße und des „Wie“ (des Umfangs) selbiger, mittels derer Geldbußen ausgeschlossen bzw. eingeschränkt werden kann.[807] Der deutsche Gesetzgeber hat hiervon Gebrauch gemacht, § 43 Abs. 3 BDSG i. V. m. § 2 Abs. 1 BDSG schließt Geldbußen gegen Behörden und sonstige öffentliche Stellen des Bundes aus (sofern diese nicht als öffentlich-rechtliche Unternehmen am Wettbewerb teilnehmen, § 2 Abs. 5 S. 1 BDSG).[808] Auch haben die Bundesländer weitestgehend eine entsprechende Verhängung von Geldbußen ausgeschlossen.[809] Hiervon ist hingegen

---

[803] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 102f., 330; siehe in diesem Zusammenhang auch *Turbék*, Justiz und KI, Friedrich-Naumann-Stiftung für die Freiheit (Hrsg.), 2026, abrufbar unter https://shop.freiheit.org/#!/Publikation/2184 (Abruf am 09.10.2026, 15:40 Uhr), S. 15, 18.

[804] Siehe hierzu ausführlich *Caroei/Brodowski/Grabmair/Koçak/Kolbe/Olheide/Paschke/Rachut/Reuß/Sigmund/Treffers/Uth/Welpe/Wörner/Zettelmeier/Langer*, MMR 2026, 619ff.

[805] *Caroei/Brodowski/Grabmair/Koçak/Kolbe/Olheide/Paschke/Rachut/Reuß/Sigmund/Treffers/Uth/Welpe/Wörner/Zettelmeier/Langer*, MMR 2026, 619, 627.

[806] BeckOK DatenschutzR/*Quaas*, 57. Ed. 1.5.2026, DS-GVO Art. 82 vor Rn. 1.

[807] BeckOK DatenschutzR/*Holländer*, 57. Ed. 1.5.2026, DS-GVO Art. 83 Rn. 79ff.

[808] BeckOK DatenschutzR/*Holländer*, 57. Ed. 1.5.2026, DS-GVO Art. 83 Rn. 81; BeckOK DatenschutzR/*Schild*, 57. Ed. 1.8.2026, BDSG § 2 vor Rn. 1; BeckOK DatenschutzR/*Brodowski*, 57. Ed. 1.5.2026, BDSG § 43 Rn. 19ff.

[809] Mit unterschiedlichen Feinheiten: § 28 LDSG BW, Art. 22 BayDSG, § 28 BlnDSG, § 32 Abs. 3 BbgDSG, § 23 Abs. 3 BremDSGVOAG, § 24 Abs. 3 HmbDSG, § 36 Abs. 2 HDSIG, § 22 Abs. 3 DSG M-V, § 20 Abs. 5 NDSG, § 32 DSG NRW, § 24 Abs. 3 LDSG Rheinland-Pfalz, § 20 Abs. 5 S. 2 DSG SL, § 19 Abs. 3 SächsDSDG, § 31 Abs. 2 DSAG LSA, § 19 Abs. 1 LDSG SH, § 61 Abs. 4 ThürDSG.

nicht auch die Verhängung einer Buße gegen Bedienstete öffentlicher Stellen erfasst, die sich nach allgemeinen Regeln richtet.[810] Als Handlungsoption kommt daher in Betracht, das grundsätzliche Verbot der Nutzung von KI-Systemen in Form der Eingabe personenbezogener Daten in Systeme, die nicht der Kontrolle der Justiz unterstehen bzw. die datenschutzrechtlichen Anforderungen nicht wahren, breit zu kommunizieren und insbesondere auf die datenschutz- sowie dienstrechtlichen Konsequenzen aufmerksam zu machen.[811] Die Einhaltung dieser rechtlichen roten Linie kann zudem flankiert werden durch die Implementierung einer Aufsichtsbehörde auch für den Bereich der justiziellen Gerichtstätigkeit, wodurch ein weiterer wirksamer sowie ersichtlicher Fokus auf die Thematik gelenkt werden kann. Überdies sollten **rechtmäßige Alternativen geschaffen** werden, sodass Richterinnen und Richter (sowie weitere Bedienstete des öffentlichen Dienstes) bereits nicht in die Verlegenheit geraten, personenbezogene Daten in Systeme, die nicht der Kontrolle der Justiz unterstehen bzw. die datenschutzrechtlichen Anforderungen nicht wahren, einzugeben. Neben dem Angebot einer eigenen datenschutzkonformen Alternative könnte (freilich unter dem Vorbehalt der Wahrung der weiteren rechtlichen Vorgaben) ggf. auch die Nutzung anderer Systeme in Betracht kommen, wenn dieser eine vollständige Anonymisierung der Daten vorausgeht (z. B. durch informationstechnische Systeme, wobei beachtet werden muss, dass diese Anonymisierungstools ihrerseits den datenschutzrechtlichen Anforderungen entsprechen müssen).[812] Die Verdeutlichung der Sanktionsmöglichkeiten im Falle eines Verstoßes gegen die datenschutzrechtlichen Regelungen in Kombination mit dem Angebot einer rechtmäßigen Alternative verspricht eine wirksame Verhinderung von Verstößen und sollte daher – als **Handlungsempfehlung** – implementiert werden.

dd. Trainingsdaten anonymisiert oder nicht personenbezogen

Soweit dies rechtlich sowie technisch möglich und sinnvoll ist, ist die Verwendung anonymisierter Daten vorzugswürdig.[813] Hierdurch kann u. a. einer systemseitigen Verknüpfung scheinbar unproblematischer, aber kontextbezogener Informationen sowie Schwierigkeiten mit etwaigen Löschungsansprüchen (oder Rechtsstreitigkeiten über diese) vorgebeugt werden sowie die Gefahr statistischer Verzerrungen und *Bias* minimiert werden.[814]
Soweit die eingesetzten KI-Modelle Falldaten auch nach Eintritt eines Löschungsgrundes zum „Dazulernen“ für weitere, zukünftige Entscheidungen

[810] Str., für die Bundesregelung siehe BeckOK DatenschutzR/*Brodowski*, 57. Ed. 1.5.2026, BDSG § 43 Rn. 19ff. m. w. N. auch zur Gegenansicht.
[811] Instruktiv zum Thema der „Accountability“ *Bovens*, Analysing and Assessing Accountability: A Conceptual Framework, 2007, European Law Journal, 13 (2007), abrufbar unter https://doi.org/10.1111/j.1468-0386.2007.00378.x (Abruf am 22.10.2026, 23:13 Uhr), S. 447ff.
[812] Vgl. *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 330.
[813] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 98ff., 252ff., 265ff.; aus rein entwicklungstechnischer Sicht die Anonymisierung nicht befürwortend Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 49.
[814] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 98f. m. w. N.; siehe auch den Hinweis zur Reproduktion von Bias aufgrund von Trainingsdaten *Turbék*, Justiz und KI, Friedrich-Naumann-Stiftung für die Freiheit (Hrsg.), 2026, abrufbar unter https://shop.freiheit.org/#!/Publikation/2184 (Abruf am 09.10.2026, 15:40 Uhr), S. 14f.

benötigen, sollte eine (echte) flächendeckende Anonymisierung anstatt einer Löschung verankert werden.[815] Durch diese Anonymisierung kann Löschungsansprüchen entsprochen werden.[816] Durch die flächendeckende Anonymisierung werden somit datenschutzrechtliche Vorgaben gewahrt und zugleich kann mit (anonymisierten) Praxisdaten sukzessive ein breites Portfolio an Trainingsdaten erstellt werden, mithilfe derer die Systeme trainiert und verbessert werden können. Die Entwicklung von Anonymisierungstools im Kontext des Justizsystems ist bereits fortgeschritten. So entwickelt das Niedersächsische Justizministerium in Kooperation mit mehreren Partnern das Anonymisierungstool ALeKS[817], dessen flächendeckender Betrieb derzeit vom Zentralen IT-Betrieb Niedersächsische Justiz vorbereitet wird.[818] Hiermit soll der geringen Veröffentlichungsanzahl von unter 3 % aller Gerichtsentscheidungen, im Sinne des Zugangs zum Recht und insbesondere der Transparenz sowie Nachvollziehbarkeit der Gerichtspraxis, entgegengewirkt werden.[819] Das Tool erzielt in einigen Rechtsgebieten bereits eine Genauigkeit von 97 % bei der Ersetzung personenbezogener Informationen durch Synonyme.[820] Als Handlungsoption bietet sich eine Fortentwicklung des ALeKS-Tools an, die sämtliche Akteninhalte automatisiert anonymisiert, wobei freilich auch die urheberrechtliche Lage (z. B. mit Blick auf anwaltliche Schriftsätze) im Auge behalten, geklärt und ggf. angepasst werden sollte. Dies kann die Grundlage dafür bereiten, dass Systeme mit Daten trainiert werden, die diskriminierungsfrei und insbesondere frei von Bias sind.
Systeme maschinellen Lernens benötigen eine erhebliche Menge an Trainingsdaten, wobei mit steigender Trainingsdatenmenge die Effektivität der KI-Anwendung steigt und für entsprechende Mustererkennung hochstandardisierte sowie normalisierte Daten benötigt werden.[821] Hierbei ermöglicht die flächendeckende E-Akten-Einführung seit dem 01.01.2026 ein hohes Potential; als Handlungsoption vorteilhaft erscheinen die Schaffung von technischen Schnittstellen der drei E-Aktensysteme der Bundesländer sowie eines Datenpools aller Bundesländer, mit Vorhaltung der Metadaten im XJustiz-Format.[822] Juristische Daten (Schriftsätze oder auch gerichtliche

---

[815] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 93, 98ff., 252ff., 311ff., 322ff.
[816] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 93, 98ff., 252ff., 311ff., 322ff.
[817] Abkürzung für: Anonymisierungs- und Leitsatzerstellungs-Kit zur smarten Veröffentlichung von Gerichtsentscheidungen.
[818] *Niedersächsisches Justizministerium*, Künstliche Intelligenz und Automation in der Justiz, abrufbar unter https://www.mj.niedersachsen.de/startseite/kunstliche_intelligenz_und_automation_in_der_justiz/kunstliche-intelligenz-und-automation-in-der-justiz-246705.html (Abruf am 31.03.2026, 19:41 Uhr).
[819] *Bundesministerium der Justiz und für Verbraucherschutz*, Gerichtsentscheidungen durch KI leichter veröffentlichen, abrufbar unter https://www.bmjv.de/DE/themen/digitales/digitalisierung_justiz/digitalisierungsinitiative/laendervorhaben/_doc/artikel_vorhaben_14_ALeKS.html (Abruf am 31.03.2026, 19:45 Uhr).
[820] *Niedersächsisches Justizministerium*, Künstliche Intelligenz und Automation in der Justiz, abrufbar unter https://www.mj.niedersachsen.de/startseite/kunstliche_intelligenz_und_automation_in_der_justiz/kunstliche-intelligenz-und-automation-in-der-justiz-246705.html (Abruf am 31.03.2026, 19:41 Uhr).
[821] *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 117ff.
[822] *Arbeitsgruppe „Einsatz von KI und algorithmischen Systemen in der Justiz“*, Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier zur 74. Jahrestagung der Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs vom 23. bis 25. Mai 2022 in Rostock, abrufbar unter https://www.justiz.bayern.de/media/images/behoerden-und-gerichte/oberlandesgerichte/nuernberg/einsatz_von_ki_und_algorithmischen_systemen_in_der_justiz.

Entscheidungen) sind indes heterogen und von geringer Strukturierung, was ihre Eigenschaft als Trainingsgrundlage erschwert.[823] Die Nutzung von Formularen, die Daten in strukturierter sowie maschinenlesbarer Form vorhalten (vgl. § 130c ZPO), sowie eine Harmonisierung der Urteilsstruktur und (auch die Strukturierung) des Parteivortrags kann dem entgegenwirken und sollte daher in Betracht gezogen werden.[824] Auch die rechtsstaatliche und insbesondere die judikative Transparenz fördernd sollte eine gesetzliche Veröffentlichungspflicht von begründeten, gerichtlichen Entscheidungen sowie Beschlüssen eingeführt werden, die unter Einsatz informationstechnischer Systeme realisiert (Pseudonymisierung bzw. Anonymisierung [insbesondere auch von Richterinnen und Richtern[825]] umfassend) werden kann.[826] Zudem erscheint jedenfalls eine Bund/Länder-übergreifende Verständigung auf einheitliche Standards im Hinblick auf Datenformate sinnvoll.[827]
Sofern die Verwendung anonymisierter Daten nicht (ausreichend) möglich ist, sollte auf die diskriminierungssensible Auswahl und Kontrolle der Daten Wert gelegt werden. Hierfür sollte auf jeder Ebene der Technik-Architektur die Bias-Problematik adressiert werden und insbesondere auf die Datenauswahl und Modellierung des Systems, technische Einwirkungen und Systemgestaltungen gegen systemseitige Diskriminierungen und Bias, Analysen zur Bias-Bekämpfung, ein dokumentiertes Datenmanagement sowie eine breite Datenbasis mit hinreichend repräsentativen, fehlerfreien und aktuellen Daten, Wert gelegt werden. Zudem sollte eine fortlaufende Daten- und Ergebnisevaluation vorgesehen werden und beachtet werden, dass gerade der Datenschutz sowie die Antidiskriminierung mit Blick auf die Trainingsdaten einander beeinflussen können (siehe ausführlich bereits im Abschnitt C.II.2.a.bb(5)).

ee. Sicherstellung von Betroffenenrechten

Der persönliche Anwendungsbereich der – den Datenschutz adressierenden – grundgesetzlichen, der menschenrechtlichen, sowie der DSGVO-Normen ist nicht deckungsgleich.[828] Insbesondere die DSGVO erfasst grundsätzlich nur natürliche Personen und nicht juristische Personen.[829] Aufgrund des vergleichbaren

pdf (Abruf am 15.04.2026, 18:58 Uhr), S. 38, 44; Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 52f.

[823] *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 131.

[824] *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 132ff.

[825] Die Zulässigkeit bzw. Erwünschtheit eines derartigen Richterprofilings wird unterschiedlich gewürdigt, vgl. hierzu und ein derartiges ablehnend *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 143ff. m. w. N.

[826] *Reformkommission Zivilprozess der Zukunft*, Abschlussbericht der Reformkommission Zivilprozess der Zukunft im Auftrag des 3. Digitalgipfels der Justizministerinnen und Justizminister des Bundes und der Länder, abrufbar unter https://www.bmjv.de/SharedDocs/Publikationen/DE/Fachpublikationen/2025_Abschlussbereicht_Reformkomission_Zivilprozess.html (Abruf am 20.04.2026, 21:09 Uhr), S. 113ff.

[827] *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 29ff.

[828] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 93, 252ff., 265ff., 305f.

[829] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 305f.

Schutzbereichs[830] kann eine einheitliche, rechtssubjektübergreifende Regelung der datenschutzrechtlichen Aspekte, orientiert an den strengsten inhaltlichen Gewährleistungsvorgaben vorteilhaft sein und zu einem weiteren Effizienzgewinn bei der Rechtsanwendung und damit bei der richterlichen Entscheidungsfindung beitragen – sei es beispielsweise dadurch, dass der Richter oder die Richterin beim entscheidungsunterstützenden Einsatz sich nur Kenntnisse bezüglich einer Regelung aneignen muss oder dass im entscheidungsersetzenden Einsatz nur die Vorgaben einer Regelung im informationstechnischen System abgebildet werden müssen. Auch könnte das hohe Schutzniveau bei den Bürgerinnen und Bürgern zur Akzeptanz des Einsatzes von KI in der richterlichen Entscheidungsfindung beitragen.
Die Betroffenenrechte nach der DSGVO sind nicht schrankenlos gewährleistet, sondern bieten – entweder in den jeweiligen Artikeln oder allgemein nach Art. 23 DSGVO – die Möglichkeit beschränkt zu werden.[831] Eine Beschränkung nach Art. 23 DSGVO muss den Schutz eines oder mehrerer Rechtsgüter gem. Art. 23 Abs. 1 DSGVO sicherstellen.[832] Bei der vorliegenden Betrachtung des Zivilverfahrens ist insbesondere Art. 23 Abs. 1 lit. f DSGVO von Bedeutung, der die richterliche Unabhängigkeit sowie den Schutz von Gerichtsverfahren zur Regelung hat.[833] Art. 23 Abs. 1 lit. f DSGVO erfasst jedoch nur die rechtsprechende Tätigkeit der Gerichte, nicht jedoch gerichtliche Verwaltungstätigkeit.[834] Auch von Relevanz kann der Tatbestand des Art. 23 Abs. 1 lit. j DSGVO, die Durchsetzung zivilrechtlicher Ansprüche, werden, wobei die Abgrenzung zu Art. 23 Abs. 1 lit. f (oder lit. i) DSGVO nicht ohne weiteres möglich ist.[835]
Gerade auch mit Blick auf den entscheidungsersetzenden Einsatz informationstechnischer Systeme bei der richterlichen Entscheidungsfindung sind Betroffenenrechte von Bedeutung.[836] Insbesondere die Frage, ob und in welchen Fällen mittels datenschutzrechtlicher Ansprüche der Quellcode oder Trainingsdaten offenbart werden müssen, wird voraussichtlich noch nachhaltige Relevanz haben.[837] Festgehalten sei jedoch an dieser Stelle, dass nach unserem Dafürhalten grundsätzlich nicht die zugrundeliegende Struktur i. S. d. Codes zwingend aufgedeckt werden muss (siehe hierzu bereits Abschnitt C.II.4.b.cc(1)).[838] Durch eine weitestmögliche sowie klare Beschränkung der Betroffenenrechte als Handlungsoption auf das vertretbare Minimum könnte die Notwendigkeit der Programmierung bestimmter Funktionen eingespart werden. Diese Ausgestaltung würde, insbesondere mit Blick auf das Auskunftsrecht, zudem

[830] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 96, 252ff., 265ff., 302ff.
[831] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 318 m. w. N.
[832] Simitis/Hornung/Spiecker gen. Döhmann/*Dix*, 2. Aufl. 2025, DS-GVO Art. 23 Rn. 21; *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 318.
[833] BeckOK DatenschutzR/*Stender-Vorwachs/Wolff/von Ungern-Sternberg*, 57. Ed. 1.8.2026, DS-GVO Art. 23 Rn. 40; *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 318; hinsichtlich Beschränkungen im Strafverfahren ist Art. 23 Abs. 1 lit. d DSGVO lex specialis gegenüber lit. f, Ehmann/Selmayr/*Bertermann*, 3. Aufl. 2024, DS-GVO Art. 23 Rn. 9.
[834] Taeger/Gabel/*Koreng*, 5. Aufl. 2026, DS-GVO Art. 23 Rn. 39, der hierbei eine Orientierung bei der Abgrenzung der Begriffe anhand der Rechtsprechung des EuGH zu Art. 55 Abs. 3 DSGVO befürwortet; *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 318 m. w. N.
[835] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 318 ausführlich sowie m. w. N.
[836] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 310ff., 319ff.
[837] Vgl. nur die verwandte Diskussion zur Beauskunftung der Scoreformel bei Auskunfteien BeckOK DatenschutzR/*Schmidt-Wudy*, 57. Ed. 1.8.2026, DS-GVO Art. 15 Rn. 78.3f. m. w. N.
[838] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 99ff., 253, 321f. ausführlich sowie m. w. N.

Geheimhaltungsinteressen sowie das Urheberrecht als anerkannte schutzwürdige Interessen unterstreichen.[839]
Hingegen sollte der gegensätzliche Ansatz, eine weitgehende Zulassung oder Stärkung der datenschutzrechtlichen Ansprüche, als **Handlungsempfehlung** verfolgt werden. Dieser Ansatz eines hohen Datenschutzniveaus beim Einsatz informationstechnischer Systeme in der richterlichen Entscheidungsfindung kann voraussichtlich das Vertrauen erhöhen. So konnte nämlich bereits in einer Studie aus dem Jahr 2015, die in Kooperation mit der Europäischen Kommission durchgeführt wurde und einen Datensatz von 925 Teilnehmern aus dem Vereinigten Königreich Großbritannien und Nordirland zur Grundlage hat, festgestellt werden, dass ein höherer Grad an wahrgenommenem datenschutzrechtlichem Schutz mit einem höheren Maß an Vertrauen in Einrichtungen einhergeht, die mit dem Datenschutz in Verbindung stehen.[840] Ein gleichwertiges und einheitliches, hohes Datenschutzniveau ist überdies gleichzeitig erklärtes Ziel der DSGVO.[841] Dieses hohe Datenschutzniveau sollte sodann auch breitenwirksam kundgegeben und erklärt werden, da ein höherer Wissensstand über die Regulierung mit einem höheren Grad an wahrgenommenem datenschutzrechtlichem Schutz einhergeht sowie ein höherer Grad an wahrgenommenem datenschutzrechtlichem Schutz mit geringeren Bedenken hinsichtlich Datenschutzrisiken einhergeht.[842] Auch klarstellende Regelungen und eine gut sichtbare Möglichkeit zur Wahrnehmung der Datenschutzrechte versprechen vor diesem Hintergrund einen Vorteil. Zudem sollte beachtet werden, dass größer wahrgenommene Belohnungen bzw. Vorteile mit einer geringeren Präferenz für regulatorische Schutzmaßnahmen sowie einem geringeren Maß an Schutzverhalten einhergehen.[843] Vor diesem Hintergrund sollte in Betracht gezogen werden, Gratifikationen in Form von (geringeren) Gerichtskosten dem Rechtsschutzsuchenden zuteilwerden zu lassen, etwa dafür, dass dieser ein entscheidungsersetzendes informationstechnisches System bei der richterlichen Entscheidungsfindung benutzt oder seine Daten als Trainingsgrundlage für die Entwicklung und Optimierung der informationstechnischen Systeme zur Verfügung stellt.

ff. Datenverarbeitung-Informationspflichten

Den Betroffenen von Datenverarbeitungsvorgängen personenbezogener Daten gegenüber sind verschiedene Auskünfte und Informationen zu erteilen. Regelungen der DSGVO können hierbei prinzipiell für das Verständnis anderer Rechtsgrundlagen Pate stehen.[844] Auf Art. 12 DSGVO aufbauend, begründen Art. 13 DSGVO und Art. 14 DSGVO Informationspflichten des Verantwortlichen, welche Betroffenen ermöglichen sollen, ihre sonstigen Rechte bezüglich seiner personenbezogenen Daten wahrnehmen zu können.[845]

---

[839] Vgl. ErwGr. 63 S. 5 der DSGVO; im Kontext des Art. 15 Abs. 1 lit. h DSGVO vgl. auch *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 321f. m. w. N. sowie Art. 15 Abs. 4 DSGVO.
[840] *Miltgen/Smith*, Exploring information privacy regulation, risks, trust, and behavior, Information & Management 52 (2015), Nr. 6, S. 741, 743ff., 748ff.
[841] ErwGr. 10 S. 1, 2 der DSGVO.
[842] *Miltgen/Smith*, Exploring information privacy regulation, risks, trust, and behavior, Information & Management 52 (2015), Nr. 6, S. 743f., 748ff.
[843] *Miltgen/Smith*, Exploring information privacy regulation, risks, trust, and behavior, Information & Management 52 (2015), Nr. 6, S. 745.
[844] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 99ff., 252ff., 265ff., 310f., 321ff.
[845] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 310 m. w. N.

Art. 13 DSGVO sowie Art. 14 DSGVO verpflichten den Verantwortlichen, der betroffenen Person bei der Datenerhebung unterschiedliche Informationen bereitzustellen, wie insbesondere Name und Kontaktdaten des Verantwortlichen bzw. seines Vertreters (lit. a), die Kontaktdaten des Datenschutzbeauftragten (lit. b), die Datenverarbeitungszwecke sowie die Rechtsgrundlage der Verarbeitung (lit. c).[846] Hierbei regelt Art. 13 DSGVO die Informationspflichten im Falle der Erhebung der personenbezogenen Daten bei dem Betroffenen selbst, während Art. 14 DSGVO den Fall betrifft, dass die Daten nicht dem Betroffenen erhoben werden, sondern bei einem Dritten.[847] Zusätzlich zu diesen Informationspflichten statuiert Art. 15 DSGVO ein Auskunftsrecht des Betroffenen mit dem Zweck, dass Betroffene sich vor Augen führen können, welche Daten verarbeitet werden, in welchem Umfang die Verarbeitung erfolgt, zur Kontrolle des Verantwortlichen sowie zur Förderung der Transparenz der Datenverarbeitung.[848] An die Zugänglichmachung bzw. Übermittlung der Informationen bzw. Auskünfte der Art. 13–15 DSGVO stellt Art. 12 Abs. 1 DSGVO Anforderungen. Der Verantwortliche hat hiernach geeignete Maßnahmen zu treffen, die Informationen in präziser, transparenter, verständlicher und leicht zugänglicher Form in einer klaren und einfachen Sprache zu übermitteln (S. 1); die Übermittlung der Informationen kann schriftlich oder in anderer Form, gegebenenfalls auch elektronisch erfolgen (S. 2). Abzustellen ist bezüglich der Anforderungen an die Art und Weise auf den durchschnittlichen Nutzer, wobei bei verschiedenen Zielgruppen ggf. verschiedene Darstellungen erforderlich sein können.[849] Insbesondere müssen die Informationen bzw. Auskünfte in leicht zugänglicher Form erfolgen, es dürfen daher keine Zugriffshürden aufgestellt werden, sondern die Betroffenen müssen die Informationen mit den ihnen zur Verfügung stehenden Mitteln erreichen können.[850] Hinsichtlich der Form (nicht i. S. v. §§ 126 ff. BGB, sondern vielmehr als „Stil“ zu verstehen) ist der Verpflichtete grundsätzlich frei, insbesondere auch die elektronische Bereitstellung ist möglich und im alltäglichen Leben dominant.[851]
Bei der Erfüllung der Pflichten zur Information der Betroffenen sollte zunächst auf die bereits etablierten sowie gängigen Methoden zurückgegriffen werden. Dies bedeutet, dass wie auch bislang üblich, die Informationen sofort ersichtlich zugreifbar sein sollten, etwa durch die Bereitstellung entsprechender Links, mittels FAQ-Listen, kontextbezogener Pop-up-Menüs beim Ausfüllen eines Online-Formulars (was insbesondere im Rahmen des entscheidungsersetzenden Einsatzes informationstechnischer Systeme von Relevanz erscheint) sowie Chatbot-Schnittstellen, bei denen die Möglichkeiten der Aufgabenübernahme durch informationstechnische Systeme ebenfalls augenfällig sind.[852] Auch sollten die

[846] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 310.
[847] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 310.
[848] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 310f. m. w. N.
[849] Str., m. w. N. zur Gegenansicht BeckOK DatenschutzR/*Quaas*, 57. Ed. 1.5.2026, DS-GVO Art. 12 Rn. 12.
[850] BeckOK DatenschutzR/*Quaas*, 57. Ed. 1.5.2026, DS-GVO Art. 12 Rn. 16.
[851] BeckOK DatenschutzR/*Quaas*, 57. Ed. 1.5.2026, DS-GVO Art. 12 Rn. 17, 27ff.
[852] *Artikel-29-Datenschutzgruppe*, WP 260 rev.01, abrufbar unter https://www.lfd.niedersachsen.de/startseite/infothek/edsa_leitlinien/vom-edsa-bestaetigte-leitlinien-158964.html (Abruf am 16.09.2026, 23:30 Uhr), Rn. 11; siehe zu weiteren geeigneten Maßnahmen a. a. O. Rn. 40.

gängigen Begrifflichkeiten wie „Datenschutzhinweise“, „Datenschutzerklärung“ o. Ä. genutzt werden sowie die Informationen prominent bereitgestellt werden.[853]
Insgesamt ist zu empfehlen, dass eine einheitliche Informationsstrategie verfolgt wird. Die weiteren bereitzustellenden Informationen, z. B. über das „Ob“ und das „Wie“ des Einsatzes informationstechnischer Systeme bei der richterlichen Entscheidungsfindung sowie die Datenschutzinformationen sollten in möglichst einheitlicher bzw. paralleler Form – soweit sinnvoll und unter Beachtung der jeweiligen Spezifika – bereitgestellt werden. Insbesondere für den entscheidungsersetzenden Einsatz informationstechnischer Systeme ist es naheliegend, dass derartige Verfahren über Online-Plattformen geführt werden, mithin in einer digitalen Umgebung. Vor diesem Hintergrund sowie mit Blick auf die Fülle an bereitzustellenden Informationen erscheinen neben Chat-Bot-Funktionen insbesondere Mehrebenen-Datenschutzerklärungen/-hinweise sinnvoll, bei denen auf der ersten Stufe ein klarer Überblick gegeben wird, mit Verweis auf die Einzelinformationen in den weiteren Stufen.[854] Zur Schaffung weiterer niederschwelliger Zugangsmöglichkeit können die Mehrebenen-Datenschutzerklärungen/-hinweise zudem mit einem anhand von inhaltlicher Tiefe und sprachlicher Schwierigkeit gestuften Informationsmodell kombiniert werden (zu diesem Ansatz siehe auch schon unter Abschnitt C.II.4.b.aa(1)). Wie auch bei der Information der Parteien und der Öffentlichkeit über das „Ob“ und das „Wie“ des Einsatzes informationstechnischer Systeme sollte hierbei darauf geachtet werden, dass keine „Informationsüberflutung“ erfolgt. Hilfreich mit Blick auf die Darstellung der Datenschutzinformationen erscheint daher die Verknüpfung verschiedener Informationskategorien, anstelle einer Darstellung der gesamten Informationen in einem Zug.[855] Auch im Bereich des entscheidungsunterstützenden Einsatzes informationstechnischer Systeme kann dieses mehrstufige Verfahren grundsätzlich angewendet werden, indem auf den mehrstufig ausgestalteten Datenschutzkomplex etwa mittels QR-Codes oder Links in der gerichtsseitigen Korrespondenz hingewiesen wird und er daraus aufgerufen werden kann.[856] Freilich sollten zudem insbesondere die Geschäftsstellen mit Blick auf fehlende digitale Kompetenz von Betroffenen sensibilisiert sein, die Informationen auch für diese Gruppe an Betroffenen adäquat bereitstellen zu können, z. B. durch eine (physische) Zuleitung einer Ausfertigung nach telefonischer Kontaktaufnahme.[857] Weitere Überschneidungen für eine einheitliche Informationsstrategie ergeben sich insbesondere hinsichtlich der möglichst frühzeitigen Bereitstellung der Informationen, vor bzw. spätestens zum Verfahrensbeginn, die Beobachtung und ggf. Adaption der Vorgehensweise anderer Rechtskreise, die barrierefreie Bereitstellung der Informationen auf (den) Webseiten (der Justiz), in Textform sowie in graphisch oder

---

853 *Artikel-29-Datenschutzgruppe*, WP 260 rev.01, abrufbar unter https://www.lfd.niedersachsen.de/startseite/infothek/edsa_leitlinien/vom-edsa-bestaetigte-leitlinien-158964.html (Abruf am 16.09.2026, 23:30 Uhr), Rn. 24, 33.

854 *Artikel-29-Datenschutzgruppe*, WP 260 rev.01, abrufbar unter https://www.lfd.niedersachsen.de/startseite/infothek/edsa_leitlinien/vom-edsa-bestaetigte-leitlinien-158964.html (Abruf am 16.09.2026, 23:30 Uhr), Rn. 35f.

855 *Artikel-29-Datenschutzgruppe*, WP 260 rev.01, abrufbar unter https://www.lfd.niedersachsen.de/startseite/infothek/edsa_leitlinien/vom-edsa-bestaetigte-leitlinien-158964.html (Abruf am 16.09.2026, 23:30 Uhr), Rn. 35.

856 *Artikel-29-Datenschutzgruppe*, WP 260 rev.01, abrufbar unter https://www.lfd.niedersachsen.de/startseite/infothek/edsa_leitlinien/vom-edsa-bestaetigte-leitlinien-158964.html (Abruf am 16.09.2026, 23:30 Uhr), Rn. 38, 40.

857 *Artikel-29-Datenschutzgruppe*, WP 260 rev.01, abrufbar unter https://www.lfd.niedersachsen.de/startseite/infothek/edsa_leitlinien/vom-edsa-bestaetigte-leitlinien-158964.html (Abruf am 16.09.2026, 23:30 Uhr), Rn. 38, 40; vgl. vorstehend im Abschnitt C.II.4.b.aa(1); siehe zu der verwandten Thematik der Barrierefreiheit des Rechtsschutzes zudem Abschnitt C.II.8.

per Video aufbereiteter Form, die Bereitstellung der Informationen auf einem einheitlichen Justizportal zum Abruf, sowie z. B. im entscheidungsersetzenden Einsatz beim Kläger die zwingende Bestätigung der Vermittlung der Informationen, bevor das System den Verfahrensfortschritt freigibt.[858]
Insbesondere mit Blick auf die Geltendmachung des Auskunftsrechts des Art. 15 DSGVO erscheint die Einrichtung eines Datenschutz-Dashboards sinnvoll, als zentrale Anlaufstelle, in dem die relevanten Informationen gebündelt angeboten bzw. abgerufen werden können (z. B. mittels einer Button-Lösung zur Anforderung der Auskünfte nach Art. 15 DSGVO).[859] Dieses Dashboard kann durch (weitere) Maßnahmen, z. B. in Form von Pull- und Push-Hinweisen ausgestattet bzw. ergänzt werden, zur nutzerorientierten Gestaltung des Datenschutzkomplexes; Push-Hinweise stellen hierbei Angaben bedarfssynchron bereit, während Pull-Hinweise den Zugang zu den Informationen (z. B. über Lerntutorien) ermöglichen.[860]
Im Zuge der einheitlichen Informationsstrategie sollte zudem auf inhaltliche Überschneidungen im Rahmen der bereitzustellenden Informationen geachtet werden. So ergibt sich die Notwendigkeit der Information der Betroffenen (z. B. einer Partei im Gerichtsverfahren) hinsichtlich der abstrakten Funktionalität des informationstechnischen Systems neben Art. 15 Abs. 1 lit. h DSGVO – z. B. im Falle des entscheidungsersetzenden Einsatzes informationstechnischer Systeme, in dem Art. 22 DSGVO einschlägig ist – auch u. a. aus dem Rechtsstaatsprinzip nach Art. 20 Abs. 3 GG.[861] Die entsprechenden Informationen sollten daher zur Ressourcenschonung einheitlich aufbereitet werden und können in der Folge in verschiedenen Zusammenhängen adäquat bereitgestellt werden.

c. Zusammenfassung

Beim entscheidungsersetzenden und entscheidungsunterstützenden Einsatz werden personenbezogene Daten im Sinne des Datenschutzrechts verarbeitet. Rote Linien ergeben sich aus dem Recht auf informationelle Selbstbestimmung aus Art. 2 Abs. 1 i. V. m. Art. 1 Abs. 1 GG, Art. 8 EMRK, Art. 8 GRCh sowie der DSGVO. Die Datenverarbeitung bedarf einer Rechtsgrundlage, ist jedoch im Rahmen der Entscheidungsfindung grundsätzlich zulässig. Für den entscheidungsersetzenden Einsatz bedarf es auch wegen Art. 22 DSGVO einer Rechtsgrundlage. Ergänzend sind als rote Linien die Gewährleistung von Datensicherheit zu verzeichnen, dass die Eingabe personenbezogener Daten in Systeme, die nicht der Kontrolle der Justiz unterstehen bzw. die die datenschutzrechtlichen Anforderungen nicht wahren, regelmäßig unzulässig ist, der Vorrang nicht personenbezogener oder anonymisierter Trainingsdaten für das Training von KI-Systemen, die Möglichkeit der Löschung personenbezogener Daten, ohne die ein Systemeinsatz grundsätzlich unzulässig ist, sowie Informationspflichten und Berichtigungsrechte.
§ 3 BDSG bietet grundsätzlich für beide Einsatzformen eine Rechtsgrundlage. Für den entscheidungsersetzenden Einsatz bedarf es jedoch zur Ausfüllung des § 3 BDSG entsprechender Aufgabenzuordnungs- bzw. Zuständigkeitsregelungen sowie wegen

[858] Vgl. hierzu im Abschnitt C.II.4.b.aa(1).
[859] *Artikel-29-Datenschutzgruppe*, WP 260 rev.01, abrufbar unter https://www.lfd.niedersachsen.de/startseite/infothek/edsa_leitlinien/vom-edsa-bestaetigte-leitlinien-158964.html (Abruf am 16.09.2026, 23:30 Uhr), Rn. 39.
[860] *Artikel-29-Datenschutzgruppe*, WP 260 rev.01, abrufbar unter https://www.lfd.niedersachsen.de/startseite/infothek/edsa_leitlinien/vom-edsa-bestaetigte-leitlinien-158964.html (Abruf am 16.09.2026, 23:30 Uhr), Rn. 39.
[861] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 146ff., 258ff., 321ff.

des Verbots automatisierter Entscheidungsfindungen nach Art. 22 Abs. 1 DSGVO einer rechtlichen Grundlage. Als Handlungsoption kommt einerseits eine gesetzliche Gestattung nach Art. 22 Abs. 2 lit. b DSGVO mit angemessenen Schutzmaßnahmen in Betracht sowie andererseits der Rekurs auf die Einwilligung beider Parteien nach Art. 22 Abs. 2 lit. c DSGVO. Weiterhin kann eine Beschränkung von Betroffenenrechten bzw. Pflichten nach Art. 23 DSGVO ergänzend in Betracht gezogen werden. Zu empfehlen ist, den **entscheidungsersetzenden Einsatz auf eine gesetzliche Grundlage im Sinne von Art. 22 Abs. 2 lit. b DSGVO** zu stellen, da ohnehin die Schaffung einer gesetzlichen Grundlage notwendig ist und so Effizienzgewinne durch taktische Verweigerungen von Einwilligungen nicht unterlaufen werden können.
Für die Datensicherheit sind insbesondere Art. 24, 25, 32 DSGVO relevant. Hinsichtlich der nach Art. 24 DSGVO zu ergreifenden Maßnahmen existiert kein verbindlicher Mindeststandard. Eine Orientierung an Katalogen mit Referenzmaßnahmen wie dem Standard-Datenschutzmodell der Datenschutzbehörden von Bund und Ländern ist hingegen möglich und sollte erfolgen. Besonderes Augenmerk sollte auf Art. 25 DSGVO gelegt werden. Dessen Anforderungen sind bei Eigenentwicklungen bereits in der Konzeptphase und bei der Entwicklung durch einen Dritten schon bei der Auswahl dieses Dritten zu beachten. Als Maßnahmen im Sinne von Art. 25 DSGVO kommen insbesondere interdisziplinäre Projektteams sowie die stringente Durchsetzung von Maßnahmen bei Datenschutzverstößen in Betracht. Da die Aufsichtsbehörden für den Datenschutz nicht für die Aufsicht über Verarbeitungen zuständig sind, die Gerichte im Rahmen ihrer justitiellen Tätigkeit vornehmen, und eine Selbstkontrolle der Justiz bislang nicht implementiert ist, sollte der Aufbau einer die gesamte richterliche Sphäre abdeckenden Aufsichtsstruktur in Betracht gezogen werden. Die Maßnahmen im Rahmen des Art. 32 DSGVO sind insbesondere mit Blick auf die konkrete Ausgestaltung, die Ausgestaltung der Ansiedelung und die jeweilige Funktionalität des konkreten Systems zu bestimmen. Synergiepotentiale zwischen Art. 32 DSGVO und Art. 9, 15 KI-VO sowie zwischen der Datenschutz-Folgeabschätzung nach Art. 35 DSGVO und dem Risikomanagement nach der KI-VO sollten genutzt werden. Mit Blick auf die Datenschutz-Folgeabschätzung ist die Aufnahme der Verarbeitungen in eine Muss-Liste gem. Art. 35 Abs. 4 DSGVO anzuraten, die im Rahmen des Einsatzes informationstechnischer Systeme in der richterlichen Entscheidungsfindung Relevanz haben. Die Bestellung eines KI-Beauftragten erscheint mit wachsender Größe einer Institution sinnvoll. Die Übertragung dieser Aufgabe auf den Datenschutzbeauftragten erscheint naheliegend. Dies ist jedoch mit Augenmaß zu handhaben und hiervon z. B. abzusehen, wenn die übertragenen Aufgaben die freien zusätzlichen Ressourcen übersteigen.
Der breite Einsatz von KI im Justizsystem und hierbei insbesondere von ChatGPT ist bereits jetzt Tatsache. Der Eingabe personenbezogener Daten in Systeme, die nicht der Kontrolle der Justiz unterstehen bzw. die die datenschutzrechtlichen Anforderungen nicht wahren, sollte dadurch begegnet werden, dass diese rote Linie und die Folgen eines Verstoßes deutlich kommuniziert und rechtmäßige Alternativen angeboten werden.
Als Trainingsdaten für das System sind möglichst anonymisierte bzw. nicht personenbezogene Daten zu verwenden. Sofern KI-Modelle Falldaten auch nach Eintritt eines Löschungsgrundes benötigen, sollte eine echte flächendeckende Anonymisierung anstatt einer Löschung erfolgen. Hierfür bietet sich eine Weiterentwicklung des Tools ALeKS an, zur automatisierten Anonymisierung sämtlicher Akteninhalte. Ergänzend kommen insbesondere die Schaffung von

technischen Schnittstellen der drei E-Aktensysteme der Bundesländer, ein länderübergreifender Datenpool, eine stärkere Harmonisierung der Struktur des Parteivortrags und von Entscheidungen in Betracht. Es sollte eine gesetzliche Pflicht zur Veröffentlichung begründeter gerichtlicher Entscheidungen eingeführt werden. Sollte die Verwendung anonymisierter Daten zu Trainingszwecken nicht (ausreichend) möglich sein, sollte auf die diskriminierungssensible Auswahl und Kontrolle der Daten Wert gelegt werden.
Mit Blick auf die Sicherstellung der Betroffenenrechte ist nicht die (mögliche) Beschränkung dieser auf ein vertretbares Minimum, sondern vielmehr die weitgehende Zulassung oder Stärkung der datenschutzrechtlichen Ansprüche zu empfehlen, vor dem Hintergrund von Vertrauensgesichtspunkten. Auch kann eine einheitliche, rechtssubjektübergreifende Regelung der datenschutzrechtlichen Aspekte, orientiert an den strengsten inhaltlichen Gewährleistungsvorgaben vorteilhaft sein, ebenso wie die Ermöglichung von Gratifikationen, beispielsweise im Form von geringeren Gerichtskosten bei Nutzung entscheidungsersetzender informationstechnischer Systeme.
Mit Blick auf die Datenverarbeitungs-Informationspflichten ist das Verfolgen einer einheitlichen Informationsstrategie zu empfehlen. Bereitzustellende Informationen, z. B. auch über das „Ob“ und das „Wie“ des Einsatzes informationstechnischer Systeme bei der richterlichen Entscheidungsfindung, sollten in möglichst einheitlicher bzw. paralleler Form zur Verfügung gestellt werden. Zudem erscheinen insbesondere mehrstufige Informationsmodelle vorteilhaft sowie die Einrichtung eines Datenschutz-Dashboards als zentrale Anlaufstelle, in dem die relevanten Informationen gebündelt angeboten bzw. abgerufen werden können.

6. Notwendigkeit einer Rechtsgrundlage

In dem ersten Arbeitspaket (Milestone 1) des Forschungsprojekts „Künstliche Intelligenz und richterliche Entscheidungsfindung“[862] sind weitere, wesentliche rote Linien herausgearbeitet worden, die sich mit dem Erfordernis des Vorliegens einer (gesetzlichen) Rechtsgrundlage für den Einsatz von KI bei der richterlichen Entscheidungsfindung befassen. Diese roten Linien sind im Wesentlichen übergreifender Natur, zeigen jedoch je nach Einsatzszenario unterschiedliche Ausprägungen.

a. Rote Linien für einen entscheidungsersetzenden Einsatz sowie Handlungsleitlinien

aa. Rote Linien entscheidungsersetzender Einsatz

In dem vorstehenden Arbeitsbericht ist für den entscheidungs*ersetzenden* Einsatz informationstechnischer Systeme herausgearbeitet worden, dass für das „Ob“ und „Wie“ eines entscheidungsersetzenden Einsatzes eine umfassende **Rechtsgrundlage** notwendig ist, die die konkreten Einsatzbereiche und -modalitäten des informationstechnischen Systems festlegt.[863]
Dies ergibt sich insbesondere aus dem Demokratieprinzip unter dem Gesichtspunkt sachlicher Legitimation[864] und dem Rechtsstaatsprinzip unter dem Gesichtspunkt des Vorbehalts des Gesetzes in Verbindung mit dem Wesentlichkeitsgrundsatz.[865]
Zudem gebietet das Recht auf den gesetzlichen Richter gemäß Art. 101 Abs. 1 Satz 2 GG und Art. 6 Abs. 1 EMRK eine gesetzliche Regelung, die festlegt, welche menschlichen Richter und Richterinnen das ordnungsgemäße Funktionieren des Systems kontrollieren sowie notfalls korrigieren und bei einer ggf. vorgesehenen Aussteuerung zuständig wären.[866] Diese Zuständigkeit muss im Vorhinein feststehen; existiert hierfür eine abstrakt-generelle Zuständigkeitsregelung, kann die konkrete Einzelzuweisung auch anhand eines Geschäftsverteilungsplanes erfolgen.[867] Auch ist einer unzulässigen Einflussnahme auf die Zuständigkeit(sordnung) des informationstechnischen Systems vorzubeugen.[868]
Zur Überwindung des Verbots einer automatisierten Entscheidung nach Art. 22 Abs. 1 DSGVO ist darüber hinaus eine spezielle Rechtsgrundlage für die Verarbeitung personenbezogener Daten gemäß Art. 22 Abs. 2 lit. b) DSGVO notwendig (bzw. Art. 23 DSGVO, relevant hier insbesondere Abs. 1 lit. f und j).[869]

---

862 Siehe hierzu ausführlich *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025.
863 *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 97ff., 126f. 142f., 217ff., 241ff., 247, 252ff., 265ff.
864 *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 125ff.
865 *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 142f.
866 *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 217ff., 247.
867 *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 217ff., 247.
868 *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 217ff., 247.
869 *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 323ff.

bb. Handlungsoptionen und -empfehlungen

(1) „Ob“ und „Wie“ des entscheidungsersetzenden Einsatzes

Mit Blick auf die Ausgestaltung einer betreffenden Rechtsgrundlage ist zu beachten, dass der entscheidungsersetzende Einsatz primär in zwei alternativen Arten implementiert werden kann: Erstens kann ein Ansatz verfolgt werden, der das Verfahren in eine dem Rechtsprechungsbegriff unterfallende (vollwertige gerichtliche) Instanz einbettet, d. h. das System trifft selbst eine letztverbindliche Entscheidung.[870] Die damit einhergehenden strengeren und damit von der anderen primären Ausgestaltungsalternative abweichenden Anforderungen[871] müssen freilich Widerhall in der entsprechenden Rechtsgrundlage und den diese begleitenden Regelungskomplexen finden und die konkreten Einsatzbereiche, die Anforderungen an ein betreffendes System, die konkreten Einsatzvoraussetzungen, die konkrete Verfahrensgestaltung inklusive etwaiger Aussteuerungsmöglichkeiten sowie die jeweilige Ansiedelung an einem bestimmten Gericht ausweisen. Zweitens kann auch ein entscheidungsersetzender Einsatz in Anlehnung an das Mahnverfahren gem. §§ 688 ff. ZPO implementiert werden.[872] Die den Einsatz entscheidungsersetzender informationstechnischer Systeme ermöglichende Rechtsgrundlage sollte hierbei die wesentlichen Strukturentscheidungen für den Einsatz explizit abbilden. Die maßgeblichen rechtlichen Rahmenbedingungen hierfür können dem vorangegangenen Milestone 1 des vorliegenden Forschungsprojekts[873] bzw. den roten Linien des vorliegenden Milestones entnommen werden.
Die weitere inhaltliche Spezifikation kann anhand der aufgezeigten Handlungsoptionen und Handlungsempfehlungen, die in praxi ergriffen werden (sollen), vorgenommen werden. Hierbei bietet es sich an, insbesondere vor dem Hintergrund der iterativen Bearbeitung sowie Überarbeitung des Komplexes des entscheidungsersetzenden Einsatzes informationstechnischer Systeme und der rasanten Entwicklung der technologischen Möglichkeiten, in der Rechtsgrundlage eine **Ermächtigung zum Rechtsverordnungserlass** – etwa durch das für die Justiz zuständige Ministerium – zu verankern. Durch den Rechtsverordnungserlass besteht die Handlungsoption, allein die notwendig einer parlamentsgesetzlichen Grundlage zuzuordnenden Aspekte mittels eines förmlichen Gesetzes zu regeln und alle weiteren Aspekte im Rahmen einer Rechtsverordnung festzulegen.[874] Indem Rechtsverordnungen schneller angepasst werden können, sind diese besser geeignet, mit rasant ändernden Realitäten regelungstechnisch Schritt zu halten, als Parlamentsgesetze mit den für sie zu durchlaufenden verfassungsrechtlich ausdifferenzierten Gesetzgebungsverfahren. Vor diesem Hintergrund bietet es sich zudem an, die Rechtsgrundlagen, insbesondere die **parlamentsgesetzliche Rechtsgrundlage, technologieoffen zu formulieren**.
Angesichts der derzeitigen konkreten Äußerung dahingehend, dass ein entscheidungs*ersetzender* Einsatz politisch nicht erstrebt wird,[875] beschränken sich

[870] Vgl. *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 116f.
[871] Vgl. *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 116f.
[872] Vgl. *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 116f.
[873] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025.
[874] Siehe zu der hierbei relevanten Wesentlichkeitstheorie Dürig/Herzog/Scholz/*Grzeszick*, 111. EL Juni 2026, GG Art. 20 Rn. 105f.
[875] Ein dem Instanzenzug vorgeschaltetes, an das Mahnverfahren angelehntes automatisiertes Verfahren ablehnend etwa vgl. *Reformkommission Zivilprozess der Zukunft*, Abschlussbericht der

die diesbezüglichen Ausführungen in diesem Bericht daher unter Beachtung der roten Linien zur Rechtsgrundlage auf grundlegende Aspekte. Gleichwohl scheinen erste klein(st)e Schritte von der vollständigen Ablehnung weg verzeichnet werden zu können.[876] Dieser Weg sollte weiter verfolgt werden und die Initiative ergriffen werden, vor dem Hintergrund der bereits derzeit bestehenden rechtlichen Möglichkeiten in hinreichend einfach gelagerten Fällen, in denen sich überschaubare Tatsachen- und Rechtsfragen von begrenzter Komplexität stellen sowie nur geringe Auslegungsspielräume bestehen und deswegen der Weg, auf dem das informationstechnische System zu seiner Entscheidung gelangt, einer mathematisch-logischen Rechenoperation gleichkommt,[877] automatisierbare Verfahren auch tatsächlich zu automatisieren und durch informationstechnische Systeme im entscheidungsersetzenden Einsatz abarbeiten zu lassen. Jedenfalls in derartigen Konstellationen, etwa einigen Fluggastrechteverfahren (z. B. Ansprüche auf Ausgleichszahlung bei Nichtbeförderung i. S. d. Art. 4 Abs. 3 i. V. m. Art. 7 Abs. 1 Fluggastrechte-VO), besteht schon jetzt die Möglichkeit zur Entlastung bzw. Freisetzung gebundener Justizressourcen.[878] Gerade in diesen Fällen stellt sich vielmehr die umgekehrte Frage, warum diese – freilich in unserem Rechtssystem eher atypischen – Fälle überhaupt (noch) das Potential der Richterinnen und Richter binden bzw. weiter binden sollten. Entsprechend sollten auch die gesetzlichen Möglichkeiten für den entscheidungsersetzenden Einsatz geschaffen und in der Folge die Entlastungsmöglichkeiten für die wichtige Ressource Justiz auch tatsächlich in der Praxis genutzt werden oder ein kritischer Blick auf die Notwendigkeit des Fortbestehens derartiger Rechtsmaterien gelegt werden.

(2) Zuständigkeitsregelungen

Neben der Ausgestaltung äquivalent zum Mahnverfahren kann der entscheidungsersetzende Einsatz informationstechnischer Systeme als **vollwertige „0.-“ (mit einem viergliedrigen Instanzenzug) oder erste Instanz (mit einem dreigliedrigen Instanzenzug)** ausgestaltet werden. Mit der Ermöglichung des

---

Reformkommission Zivilprozess der Zukunft im Auftrag des 3. Digitalgipfels der Justizministerinnen und Justizminister des Bundes und der Länder, abrufbar unter https://www.bmjv.de/SharedDocs/Publikationen/DE/Fachpublikationen/2025_Abschlussbereicht_Reformkomission_Zivilprozess.html (Abruf am 20.04.2026, 21:09 Uhr), S. 99ff.; ein gerichtliches Vorentscheidungsverfahren in Zivilsachen ablehnend. Ein automatisiertes Vorentscheidungsverfahren (als Art erweitertes Mahnverfahren verstehend), aber auch in der Art einer gerichtlichen Vorentscheidung wird als im Widerspruch zum zentralen Angebot des Staates und der Justiz zur Konfliktlösung durch (menschliche) Richterinnen und Richter nach rechtsstaatlichen Maßstäben abgelehnt durch die *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, Münchener Thesen zum Zivilprozess der Zukunft, abrufbar unter https://www.olg-duesseldorf.nrw.de/infos/Zivilprozess-der-Zukunft/Muenchener-Thesen-zum-Zivilprozess-der-Zukunft.pdf (Abruf am 26.02.2026, 18:27 Uhr), S. 7, These A.5; *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 4.

[876] Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 82f.

[877] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 87ff., 106ff., 115ff., 143ff., 170ff., 229ff., 240ff., 255f., (258ff.,) 265ff., 334f., 338f., 344, 356ff.

[878] Vgl. nur *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 88.

entscheidungsersetzenden Einsatzes ist eine gesetzliche Regelung zur Aussteuerung sowie eine Anpassung der sachlichen, örtlichen sowie funktionellen Zuständigkeitsordnung vorzusehen. Hierzu bedarf es einer Anpassung des GVG und der ZPO, die ebenfalls auf die Besonderheiten des Art. 22 DSGVO eingeht und dort explizit die Verarbeitung personenbezogener Daten gestattet (siehe hierzu bereits zuvor im Abschnitt C.II.5.b.aa).

Bei der **Anpassung der sachlichen Zuständigkeit** für den Einsatz entscheidungsersetzender informationstechnischer Systeme besteht als Handlungsoption zunächst, sich an den bisherigen Gegebenheiten zu orientieren. Bei der Zuweisung der bürgerlich-rechtlichen Streitigkeiten im Rahmen einer automatisierten, vollwertigen gerichtlichen Eingangsinstanz an das Land- oder Amtsgericht, kann wie derzeit die Ausgestaltung gewählt werden, unter Beibehaltung der landgerichtlichen Allzuständigkeit, die amtsgerichtliche Zuständigkeit nach dem Enumerationsprinzip, sowohl streitwertabhängig als auch für bestimmte Materien streitwertunabhängig zu eröffnen.[879] Vor dem Hintergrund des naheliegenden Führens von Verfahren mittels informationstechnischer Systeme im entscheidungsersetzenden Einsatz über eine Online-Plattform[880] und mit Blick auf die Kernmaterie des Zivilprozesses als dem bürgerlichen Recht, das gem. Art. 74 Abs. 1 Nr. 1 Var. 1 GG Gegenstand der konkurrierenden Gesetzgebung ist und mittels des Bürgerlichen Gesetzbuches auf Bundesebene eine weitgehend abschließende Regelung erfahren hat[881], bietet sich auch vor dem Hintergrund, der sich durch den Einsatz informationstechnischer Systeme eröffnenden Konzentrationsmöglichkeit, eine bundesweit bei einem Bundesland konzentrierte und ggf. sogar in Bundeshand liegende gerichtliche Zentralzuständigkeit an.

Aufgrund der Spezialisierung der informationstechnischen Systeme auf bestimmte Materien im entscheidungsersetzenden Einsatz sollte diese Zuweisung streitwertunabhängig und anhand der Materien (im Einklang mit den existierenden spezialisierten entscheidungsersetzenden System[modul]en) erfolgen. Sie kann sukzessive angepasst sowie ergänzt werden. Eine bundesweit exklusive Zuweisung für entscheidungsersetzende Systeme kann förderlich mit Blick auf kontrollierbare Betriebsumgebungen und die Datenhoheit wirken und garantiert von sich aus bundesweit einheitliche Standards; zudem werden dadurch Insellösungen und Parallelentwicklungen vermieden, was auch mit Blick auf Ressourcen und einheitliche Prioritätssetzung vorteilhaft wirkt.[882]

---

[879] MüKoZPO/*Pabst*, 6. Aufl. 2022, GVG § 23 Rn. 1; Musielak/Voit/*Wittschier*, 23. Aufl. 2026, GVG § 23 Rn. 1; die Einführung eines beschleunigten Online-Verfahrens zunächst bis zu einem Streitwert von 5000 € befürwortend Thesenpapier der Arbeitsgruppe „Modernisierung des Zivilprozesses“ im Auftrag der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs,* abrufbar unter https://www.justiz.bayern.de/media/images/behoerden-und-gerichte/oberlandesgerichte/nuernberg/thesenpapier_der_arbeitsgruppe.pdf (Abruf am 26.02.2026, 13:01 Uhr).

[880] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 357 m. w. N.

[881] BeckOK GG/*Seiler*, 66. Ed. 15.8.2026, GG Art. 74 Rn. 2.1.

[882] Diese Ausgestaltung entspricht damit der dritten und vierten Handlungsempfehlung des Beschlusses der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs* im Mai 2026 zum Grundlagenpapier 2026 zum Einsatz von KI und algorithmischen Systemen in der Justiz, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 20:49 Uhr); vgl. auch *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 32ff., wo eine

Der Fall einer ggf. notwendigen **Aussteuerung einer Streitigkeit** vom entscheidungsersetzenden informationstechnischen System **an einen menschlichen Richter bzw. eine menschliche Richterin** entspricht mit Blick auf die sachliche Zuständigkeit materiell der derzeitigen Rechtslage der Zuweisung einer bestimmten Streitmaterie. Im Rahmen der Zuweisung der sachlichen Zuständigkeit im Aussteuerungsfalle kann sich daher an der derzeitigen Rechts(- und Diskussions)lage orientiert werden.
Auch die **Anpassung der Regelungen zur örtlichen Zuständigkeit**[883] des Eingangsgerichts für den Einsatz entscheidungsersetzender informationstechnischer Systeme ist zu betrachten. Die Regelungen zur örtlichen Zuständigkeit sind primär mit Blick auf die am Rechtsstreit beteiligten Parteien sowie die Art der Streitsache gefasst.[884] Die naheliegenderweise im entscheidungsersetzenden Einsatz informationstechnischer Systeme (über eine) Online(-Plattform) ablaufenden Verfahren können bei bestimmten Gerichten – z. B. nach Verfahrensart oder (vorzugswürdig) nach Streitmaterie – zentral angesiedelt werden.[885] Dies ist (und daher als Vergleichsobjekt naheliegend) beispielsweise in Nordrhein-Westfalen für das weitestgehend automatisierte Mahnverfahren vorgesehen – gemäß § 6 JuZuVO NRW[886] sind Mahnverfahren aus den Bezirken der Amtsgerichte in den Oberlandesgerichtsbezirken Düsseldorf und Hamm dem Amtsgericht Hagen zugewiesen, die entsprechenden Mahnverfahren aus dem Oberlandesgerichtsbezirk Köln dem Amtsgericht Euskirchen; als Service-Rechenzentrum wird von beiden Amtsgerichten der Landesbetrieb IT.NRW mit dem Standort Hagen verwendet.[887] Die initiale, automatisierte Behandlung von Verfahren könnte daher konzentriert bestimmten Amts- bzw. Landgerichten zugewiesen werden. Derartige Zuständigkeitskonzentrationen könnten bei entsprechender Verständigung auch Bundesländer-übergreifend vorgenommen werden.[888] Bei einer bundesweit konzentrierten bzw. ggf. sogar in Bundeshand liegenden gerichtlichen Zentralzuständigkeit erübrigt sich freilich die Frage der örtlichen Zuständigkeit in Bezug auf die Ansiedelung des informationstechnischen Systems selbst – abgesehen

KI-Plattform in einem einheitlichen KI-Ökosystem der Justiz vorgeschlagen wird, um Insellösungen und Parallelentwicklungen zu vermeiden. Hierdurch sollen insbesondere Synergien sowie die Mit- bzw. Nachnutzungen von unterstützenden KI-Systemen ermöglicht werden.

[883] Hierzu siehe *Adolphsen*, Zivilprozessrecht, 9. Aufl. 2025, § 6 Rn. 40ff.

[884] *Adolphsen*, Zivilprozessrecht, 9. Aufl. 2025, § 6 Rn. 41.

[885] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 357 m. w. N.; Thesenpapier der Arbeitsgruppe „Modernisierung des Zivilprozesses“ im Auftrag der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://www.justiz.bayern.de/media/images/behoerden-und-gerichte/oberlandesgerichte/nuernberg/thesenpapier_der_arbeitsgruppe.pdf (Abruf am 26.02.2026, 13:01 Uhr); für den Bereich eines beschleunigten Verfahrens, das von menschlichen Richtern geführt wird ablehnend, mit dem Argument, dass die gesamte Justiz einen Digitalisierungsschub benötige, indes eine Zentralisierung nach Streitmaterien empfehlend *Meller-Hannich*, AnwBl 2021, 288, 289.

[886] Justizzuständigkeitsverordnung vom 4. Dezember 2024 (GV. NRW. S. 1144), die zuletzt durch Verordnung vom 5. Dezember 2025 (GV. NRW. S. 1152) geändert worden ist.

[887] *Amtsgericht Euskirchen*, Grundsätzliches und Zuständigkeit, abrufbar unter https://www.ag-euskirchen.nrw.de/behoerde/zema/grundsaetzliches-zustaendigkeit/index.php (Abruf am 26.02.2026, 15:53 Uhr).

[888] Die mehrere Bundesländer umfassende Zusammenlegung von Gerichten wurde auch bereits in der Vergangenheit erwogen, vgl. *dpa/sh/LTO-Redaktion*, Ost-Länder wollen Gerichte zusammenlegen, abrufbar unter https://www.lto.de/recht/nachrichten/n/justiz-ost-laender-wollen-gerichte-zusammenlegen (Abruf am 18.03.2026, 12:36 Uhr).

von verschiedenen Standorten des Gerichts selber.[889] Auch könnten die Rechtsmittelinstanzen konzentriert bestimmten Gerichten zugewiesen werden.
Die Konzentration von Verfahren bei bestimmten Gerichten kann die organisatorische sowie inhaltliche Spezialisierung dieser Gerichte für die entsprechenden Bereiche und damit die Verfahrenseffizienz und Qualität der Entscheidung erhöhen.[890] Bei einer Aussteuerung des Verfahrens an menschliche Richterinnen und Richter könnte – dem Konzentrationsgedanken folgend –, das Verfahren an die Gerichte ausgesteuert werden, an denen die automatisierten Verfahren selbst erfolgen. Mit Blick auf den Aspekt der Verfahrensbeteiligten[891], könnte das Verfahren als Handlungsoption jedoch auch an die Gerichte ausgesteuert werden, die ansonsten (d. h. läge eine Vorschrift nicht vor, die die Zuständigkeit konzentrieren würde) zuständig wären, über die Streitigkeit zu befinden. Dies erscheint insbesondere mit Blick auf den bei der örtlichen Zuständigkeit wesentlichen Aspekt der am Rechtsstreit Beteiligten[892] vorteilhaft, als diese dann keinen unverhältnismäßig langen Weg zum Gericht auf sich nehmen müssen. Ebenso erscheint dies vor dem Hintergrund sinnvoll, als an diesen Gerichten derzeit bereits in die Materie eingearbeitete Richterinnen und Richter arbeiten sowie zudem auch perspektivisch unproblematisch, als die Kompetenz in den jeweiligen Themenbereichen zur Behandlung dieser Fallgestaltungen weiterhin hinreichend an den einzelnen Gerichten verbleiben wird. So können nämlich zwar – das Beispiel der Fluggastrechteverfahren aufgreifend – etwa Ansprüche auf Ausgleichszahlung bei Nichtbeförderung gem. Art. 4 Abs. 3 i. V. m. Art. 7 Abs. 1 Fluggastrechte-VO als hinreichend einfach gelagerte Fälle automatisiert entschieden werden, indes werden auch weiterhin andere – jedenfalls derzeit – nicht-automatisierbare fluggastrechtliche Verfahren von menschlichen Richterinnen und Richtern an den Gerichten verhandelt werden.
Neben der vorgenannten Aussteuerung an die ansonsten zuständigen Gerichte besteht als weitere Handlungsoption eine – auch Bundesland-übergreifende – Konzentration der örtlichen Zuständigkeit im Aussteuerungsfalle. Hierbei bieten sich insbesondere gut vernetzte Knotenpunkte im Hoheitsgebiet der Bundesrepublik Deutschland an, damit eine verhältnismäßig gleichschwellige Erreichbarkeit des Gerichts für die Verfahrensbeteiligten gewährleistet werden kann. Augenfällige Optionen hierfür sind beispielsweise die Städte Frankfurt am Main sowie München, Köln bzw. Düsseldorf, Hamburg und Berlin mit ihrer ausgebauten Vernetzung auch durch den Bahn- und Flugverkehr. Anzahl und Verortung wären bei Wahl dieser Option freilich maßgeblich vom politischen (Kooperations)Willen abhängig.
Aufgrund der Notwendigkeit eines Online-Verfahrens in Verbindung mit Barrierefreiheitsmaßnahmen bezüglich des Rechtsschutzes, wie **physischen Rechtsantragstellen** (siehe hierzu Barrierefreiheit des Rechtsschutzes, Abschnitt C.II.8), die an den (bzw. ausgewählten) Zivilgerichten implementiert werden, kann

[889] So hat beispielsweise der Bundesgerichtshof Dienststellen in Karlsruhe sowie in Leipzig, *Bundesgerichtshof*, Anfahrt zum Bundesgerichtshof in Karlsruhe, abrufbar unter https://www.bundesgerichtshof.de/DE/Service/Kontakt/AnfahrtKarlsruhe/anfahrtKarlsruhe_node.html; *Bundesgerichtshof*, Anfahrt zur Außenstelle des Bundesgerichtshofs in Karlsruhe, abrufbar unter https://www.bundesgerichtshof.de/DE/Service/Kontakt/AnfahrtKarlsruheRQ/anfahrtKarlsruheRQ_node.html; *Bundesgerichtshof*, Anfahrt zur Dienststelle Leipzig des Bundesgerichtshofs, abrufbar unter https://www.bundesgerichtshof.de/DE/Service/Kontakt/AnfahrtLeipzig/anfahrtLeipzig_node.html, (jeweils Abruf am 19.10.2026, 16:59 Uhr).
[890] Vgl. die Begründung zu § 3 des Verbraucherrechtedurchsetzungsgesetzes, Skauradszun/*Gödicke*, 1. Aufl. 2024, VDuG § 3 Rn. 10 mit Verweis auf BT-Drucks. 20/6520, S. 71.
[891] *Adolphsen*, Zivilprozessrecht, 9. Aufl. 2025, § 6 Rn. 41.
[892] *Adolphsen*, Zivilprozessrecht, 9. Aufl. 2025, § 6 Rn. 41.

auch Bedenken[893] mit Blick auf den Zugang zum Recht in der Fläche wirksam begegnet werden. Größere (zentralisierte) Standorte versprechen zudem Vorteile mit Blick auf die Förderung der Digitalisierung; vor Ort am Gericht bestehende menschliche IT-Ressourcen haben sich bereits als vorteilhaft in der Praxis erwiesen.[894] Auch können Vorteile mit Blick auf die Effektivität der Organisation sowie bei der Vereinfachung der Personalsteuerung erwartet werden.[895] Zudem kann als Vorteil die Resilienz gegenüber Entwicklungen wie dem demographischen Wandel genannt werden.[896]

Mit Blick auf die **funktionelle Zuständigkeit**[897] sind Regelungen zu schaffen, die den Einsatz informationstechnischer Systeme im entscheidungsersetzenden Einsatz betreffen. Dies betrifft maßgeblich die **Verteilung innerhalb eines Gerichts sowie innerhalb der Gerichtsorgane**.[898] Mit Blick auf die Verteilung der Aufgabe innerhalb eines Gerichts kann die Zuteilung des entscheidungsersetzenden informationstechnischen Systems entsprechend der von ihm inhaltlich wahrgenommenen Aufgabe der jeweiligen Gerichtssparte vorgenommen werden (Bsp.: Bei einer Aufgabenwahrnehmung in Vollstreckungssachen etwa dem Amtsgericht als Vollstreckungsgericht oder bei einer Aufgabenwahrnehmung in Prozesssachen dem Amtsgericht als Prozessgericht). Dies betrifft sowohl die Ausgestaltung des Einsatzes des informationstechnischen Systems als vollwertige gerichtliche Instanz als auch die Ausgestaltung als Mahnverfahrensäquivalent. Innerhalb des gerichtlichen Geschäftsverteilungsplans können die näheren Modalitäten festgelegt werden, nach denen entscheidungsersetzende informationstechnische Systeme im konkreten Fall zur Entscheidung berufen sind. Naheliegend erscheint die Zuteilung von konkreten Sachgebieten[899], für die das entscheidungsersetzende informationstechnische System geeignet ist.[900] Fragen der Aufgabenverteilung innerhalb eines Spruchkörpers erübrigen sich hingegen mit Blick auf den entscheidungsersetzenden Einsatz informationstechnischer Systeme selbst.

---

[893] Vgl. im Fall der Zusammenlegung der Arbeitsgerichte im Land Nordrhein-Westfalen, *Amos*, NRW will Arbeitsgerichte zusammenstutzen, abrufbar unter https://rsw.beck.de/aktuell/daily/meldung/detail/diskussionspapier-nrw-arbeitsgerichte-strukturreform-zusammenlegung (Abruf am 18.03.2026, 14:59 Uhr).

[894] Vgl. *Ministerium der Justiz des Landes Nordrhein-Westfalen,* Arbeitsgerichtsbarkeit der Zukunft in Nordrhein-Westfalen Diskussionspapier für den weiteren Beteiligungsprozess, abrufbar unter https://www.justiz.nrw.de/sites/default/files/2025-11/2025-11-11%20Diskussionspapier.pdf (Abruf am 18.03.2026, 15:28 Uhr) S. 5, 7.

[895] Vgl. *Ministerium der Justiz des Landes Nordrhein-Westfalen,* Arbeitsgerichtsbarkeit der Zukunft in Nordrhein-Westfalen Diskussionspapier für den weiteren Beteiligungsprozess, abrufbar unter https://www.justiz.nrw.de/sites/default/files/2025-11/2025-11-11%20Diskussionspapier.pdf (Abruf am 18.03.2026, 15:28 Uhr) S. 5f.

[896] Vgl. *Ministerium der Justiz des Landes Nordrhein-Westfalen,* Arbeitsgerichtsbarkeit der Zukunft in Nordrhein-Westfalen Diskussionspapier für den weiteren Beteiligungsprozess, abrufbar unter https://www.justiz.nrw.de/sites/default/files/2025-11/2025-11-11%20Diskussionspapier.pdf (Abruf am 18.03.2026, 15:28 Uhr) S. 6.

[897] Zu der Begrifflichkeit der funktionellen Zuständigkeit Stein/*Roth*, 24. Aufl. 2024, ZPO § 1 Rn. 58ff.

[898] Stein/*Roth*, 24. Aufl. 2024, ZPO § 1 Rn. 58; hinsichtlich der Verteilung auf verschiedene Gerichte (Instanzenzug) siehe Abschnitt C.II.7.b.

[899] Saenger/*Rathmann*, 10. Aufl. 2023, GVG § 21e Rn. 7. Denkbar wäre z. B. auch eine Kombination aus Sachgebiet und Anfangsbuchstaben des Beklagten-Namens, sofern eine parallele Bearbeitung des Aufgabengebiets sowohl von menschlichen Richterinnen und Richtern als auch von entscheidungsersetzenden informationstechnischen Systemen präferiert wird. So könnte die eine Buchstabenreihe in diesem Falle vom informationstechnischen System und die andere Buchstabenreihe menschlich gerichtet werden. Hierfür könnte u. a. sprechen, einer vollständigen Standardisierung eines Rechtsprechungssachgebiets entgegenzuwirken.

[900] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 217ff.

Maßgebliches Augenmerk ist hinsichtlich der funktionellen Zuständigkeit beim Einsatz eines entscheidungsersetzenden informationstechnischen Systems auf den **Aussteuerungsfall** zu legen. Bezüglich dieses Falles bietet es sich an, mit Blick auf die Aufgabenverteilung bei Gericht nach Art der Aufgabe, zwischen mehreren gleichartigen Spruchkörpern, sowie innerhalb der einzelnen Spruchkörper, den ausgesteuerten Fall so zuzuteilen, als wäre kein Einsatz eines entscheidungsersetzenden informationstechnischen Systems vorgesehen gewesen. Auch eine Verteilung an mit sachnahen Aufgaben betraute Spruchkörper kann vorgesehen werden. In beiden Fällen würde so gesichert sein, dass in die Materie ohnehin bereits inhaltlich eingearbeitete Richterinnen und Richter mit der Aufgabe betraut werden, mithin eine sachgerechte Aufgabenwahrnehmung. Die bereits existierenden Regelungen bieten zudem einen tauglichen Anknüpfungspunkt zur Aufgabenverteilung innerhalb eines Spruchkörpers im Aussteuerungsfalle. Gem. § 22 Abs. 1, 4 GVG agieren an den Amtsgerichten in Zivilsachen stets Richter als Einzelrichter.[901] Die mit drei Richtern ausgestatteten Kammern am Landgericht, §§ 60, 75 GVG, entscheiden regelmäßig in Form des originären Einzelrichters, §§ 348 Abs. 1 S. 1 ZPO bzw. bei Nichtvorliegen der Voraussetzungen, nach Kammerbeschluss, in einfach gelagerten Fällen nach § 348a ZPO als sog. obligatorischer Einzelrichter.[902] Die Abgrenzung, ob eine landgerichtliche Zivilsache durch einen Einzelrichter entschieden wird oder durch die Kammer, ist insbesondere von der Schwierigkeit und Bedeutung des Falles abhängig, vgl. §§ 348 Abs. 3 Nr. 1 u. 2, 348a Abs. 1 Nr. 1 u. 2, Abs. 2 Nr. 1 ZPO. Die Abgrenzung für die funktionelle Zuständigkeit innerhalb der Landgerichte nach Schwierigkeit und Bedeutung ist auch für die nach derzeitigem Stand automatisierbaren (und insoweit Aussteuerungen zugänglichen) Fälle sachlich passend. Mit Blick auf den Umfang der anfallenden ausgesteuerten Fälle sowie mit Blick auf eine ggf. gewählte intensivierte örtliche Zuständigkeitskonzentration können auch eigene, überwiegend bis hin zu ausschließlich mit Aussteuerungsfällen betraute Spruchkörper (z. B. sog. „Automatisierungskammern“) geschaffen werden. Freilich ist für diese Option die inhaltliche Varianzbreite der ausgesteuerten Fälle relevant, da entscheidungsersetzende informationstechnische Systeme – bei Einhaltung der rechtlichen roten Linien – für eine Pluralität an Fallgestaltungen implementiert werden können. So erscheint die Bündelung mehrerer inhaltlich wenig verwandter Fallgestaltungen bei einem Spruchkörper unter Effizienzgesichtspunkten nur bedingt zweckmäßig.
Auch für die richterliche Aufsichts- und Kontrolltätigkeit bezüglich des informationstechnischen Systems sowie für notfalls notwendige Korrekturen sind Zuständigkeitsregelungen vorzusehen. Für die inhaltliche Ausgestaltung kommt eine Vielzahl an Möglichkeiten in Betracht (siehe hierzu insbesondere bereits Abschnitt C.II.2.a.bb). Mit Blick auf die sich anbietende Konzentration der Ansiedelung des Betriebs des entscheidungsersetzenden informationstechnischen Systems kann sich gleichsam auch die entsprechende Konzentration der Zuständigkeiten für die Kontrolle des ordnungsgemäßen Funktionierens anbieten. Auch könnte entsprechend die ggf. notwendige Korrekturtätigkeit angesiedelt werden. Eine maßgebliche Handlungsoption besteht darin, das Vorhandensein einer interdisziplinären hohen Expertise für die Aufsichts- und Kontrolltätigkeiten an sich sowie diese in hinreichender Art und Weise normativ – insbesondere auch für die funktionelle Zuständigkeit – festzulegen. Dies erscheint mit Blick auf die Wesentlichkeit des ordnungsgemäßen

[901] Siehe so auch ausdrücklich *Adolphsen*, Zivilprozessrecht, 9. Aufl. 2025, B. § 6 Rn. 25.
[902] *Adolphsen*, Zivilprozessrecht, 9. Aufl. 2025, B. § 6 Rn. 26.

Funktionierens der entscheidungsersetzenden informationstechnischen Systeme – so können entscheidungsersetzende Systeme eine Vielzahl von Fällen gleichsam „auf einen Schlag“ abarbeiten – und der damit einhergehenden Verantwortungsfülle der Aufsichts- und Kontrolltätigkeit, sinnvoll. Das Ausmaß der Ausgestaltung ist insbesondere von den für die Aufsicht und Kontrolle notwendigen Kapazitäten abhängig, ergänzt um den für die Korrekturtätigkeiten notwendigen Aufwand. Weitere relevante Faktoren für die Ausgestaltung sind insbesondere der Umfang an und die Art der verschiedenen Systeme(n) sowie der Umfang der von den Systemen wahrgenommenen, inhaltlich unterschiedlichen, Aufgaben. Hierbei ist zu beachten, dass bereits aufgrund der verschiedenen rechtlichen Fallgestaltungen präsumtiv mehrere, jeweils mit anderen Materien inhaltlich vertraute, mit (juristisch gebildeten) Richterinnen und Richter besetzte Aufsichts- und Kontrolleinheiten notwendig werden (können), um die Aufgaben wahrnehmen zu können. Hierbei könnte als Handlungsoption implementiert werden, etwa bei einer (dem Gedanken der „Automatisierungskammer“ folgenden) konzentrierten Ansiedelung der Zuständigkeit dem Grunde nach, entsprechend dem Arbeitsanfall eine/n oder mehrere Richter bzw. Richterin/nen aus verschiedenen thematisch passenden Spruchkörpern von (anderen, ohnehin bestehenden) Gerichten – unter entsprechender Freistellung – für die Aufsichts- und Kontrolltätigkeit heranzuziehen. Dies würde die konkrete Vertrautheit aus rechtlicher Sicht mit der konkreten Fallmaterie sicherstellen sowie aktuell halten. Mit Blick auf den derzeitigen Stand, nach dem der entscheidungsersetzende Einsatz informationstechnischer Systeme lediglich in hinreichend einfach gelagerten Fallkonstellationen denkbar ist, könnte die v. a. inhaltliche Überwachung auch den in den Geschäftsverteilungsplänen der dezentralen Gerichtsorganisation zuvor bestimmten Richterinnen und Richtern zugewiesen werden (siehe hierzu bereits Abschnitt C.II.2.a.bb(1)). Eine weitere Handlungsoption ist perspektivisch die vollständige Verwendung bzw. originäre Ansiedelung von Richterinnen und Richtern in den Aufsichts- und Kontrollkörpern. Dies könnte auch ergänzt werden um benannte, fachlich konkret versierte Kontaktpersonen für bestimmte Themenfelder aus der bestehenden dezentralen Gerichtsorganisation, die auf Anfrage hin beratend bzw. unterstützend tätig werden. Die Konzentration der Ansiedelung der Aufsichts- und Kontrollinstitutionen ist in verschiedenen graduellen Ausgestaltungen möglich, denkbar wären insbesondere die Ansiedelung bei mehreren höherinstanzlichen Gerichten, um die dort vorhandene vorherrschende hohe rechtliche Expertise zu nutzen, sowie eine bundeseinheitliche Lösung. Jedenfalls bei einer bundeseinheitlichen Ausgestaltung kommt auch das Schaffen eines die thematischen Themenfelder umfassend funktionell mit Aufsichts- und Kontrollkörpern abbildenden Aufsichts- und Kontroll-„Gerichts“ in Betracht (siehe hierzu auch Abschnitt C.II.2.a.bb(1)). Grundlegend erscheint als Handlungsoption vorteilhaft – unter Beachtung des Aspekts der Erforderlichkeit sowie des staatlichen Ressourceneinsatzes – eine weitestgehende Konzentration anzustreben, um Parallelentwicklungen zu vermeiden sowie Effektivitäts-, Spezialisierungs- und Personal- und Qualitätsvorteile nutzen zu können.[903] Zudem sollten Potentiale aus den

[903] Vgl. *Ministerium der Justiz des Landes Nordrhein-Westfalen,* Arbeitsgerichtsbarkeit der Zukunft in Nordrhein-Westfalen Diskussionspapier für den weiteren Beteiligungsprozess, abrufbar unter https://www.justiz.nrw.de/sites/default/files/2025-11/2025-11-11%20Diskussionspapier.pdf (Abruf am 18.03.2026, 15:28 Uhr), S. 5ff.; vgl. die Begründung zu § 3 des Verbraucherrechtedurchsetzungsgesetzes, Skauradszun/*Gödicke*, 1. Aufl. 2024, VDuG § 3 Rn. 10 mit Verweis auf BT-Drs. 20/6520, S. 71; zu mit Konzentrationen verbundenen Vorteilen vgl. auch Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts*

Fällen der Aussteuerung sowie eingelegter Rechtsmittel gegen entscheidungsersetzende informationstechnische Systeme (siehe hierzu nachfolgend, Abschnitt C.II.7.b) genutzt werden. Für diese Fälle kann als Handlungsoption die Pflicht festgelegt werden, dass die mit den Fällen (ohnehin) betrauten Richterinnen und Richter die systemseitig vorgenommenen Fehler systematisch zu protokollieren und zu übermitteln haben sowie im weiteren Verlauf dem Optimierungsprozess des Systems zur Seite stehen müssen. Auf diese Weise können die Erkenntnisse der in die konkreten Fälle eingearbeiteten (Rechtsmittel-)Richterinnen und -Richter genutzt werden und die Notwendigkeit einer anderweitigen separaten Einarbeitung in den Fall gemildert bzw. verhindert werden.

cc. Zusammenfassung

Für das „Ob“ und das „Wie“ eines **entscheidungsersetzenden Einsatzes** bedarf es nach dem Demokratie- und dem Rechtsstaatsprinzip einer **umfassenden Rechtsgrundlage**, die die konkreten Einsatzbereiche und Einsatzmodalitäten festlegt. Das Recht auf den gesetzlichen Richter verlangt zudem eine im Vorhinein feststehende Zuständigkeitsregelung hinsichtlich dessen, welche Richterinnen und Richter das ordnungsgemäße Funktionieren des Systems kontrollieren, es notfalls korrigieren und für Aussteuerungsfälle zuständig sind. Es ist einer unzulässigen Einflussnahme auf die Zuständigkeit(sordnung) vorzubeugen. Datenschutzrechtlich ist eine Rechtsgrundlage (nach Art. 22 Abs. 2 lit. b) DSGVO) erforderlich, zur Überwindung des Verbots einer automatisierten Entscheidung nach Art. 22 Abs. 1 DSGVO.
Die zu schaffende Rechtsgrundlage sollte die wesentlichen Strukturentscheidungen für den Einsatz explizit abbilden. Dies betrifft auch die Frage der Ausgestaltung als vollwertige Instanz oder ähnlich dem Mahnverfahren. Die maßgeblichen rechtlichen Rahmenbedingungen hierfür können dem vorangegangenen Milestone 1 bzw. den roten Linien des vorliegenden Berichts entnommen werden. Es bietet sich an, in der parlamentsgesetzlichen Grundlage nur die Aspekte zu regeln, die notwendig in dieser zu verorten sind, und in dieser eine Ermächtigung zum Rechtsverordnungserlass niederzulegen. In der auf dieser Grundlage erlassenen Rechtsverordnung können sodann alle weiteren Aspekte geregelt werden. Die Rechtsgrundlagen sollten technologieoffen formuliert sein. Für hinreichend einfach gelagerte Fälle (wie Ansprüche auf Ausgleichszahlung bei Nichtbeförderung nach der Fluggastrechte-VO) sollten die gesetzlichen Möglichkeiten für den entscheidungsersetzenden Einsatz geschaffen und damit die Entlastungsmöglichkeiten für die wichtige Ressource Justiz tatsächlich genutzt oder aber die Notwendigkeit des Fortbestehens derartiger Rechtsmaterien kritisch geprüft werden.
Durch Anpassung von GVG und ZPO sind die für den entscheidungsersetzenden Einsatz notwendigen Regelungen zur Aussteuerung und Anpassung der sachlichen, örtlichen und funktionellen Zuständigkeit vorzusehen. Mit Blick auf die sachliche Zuständigkeit sollte die Zuweisung streitwertunabhängig und anhand der Materien (im Einklang mit den existierenden spezialisierten entscheidungsersetzenden System[modul]en) erfolgen. Hierbei bietet sich insbesondere eine bundesweit konzentrierte Zentralzuständigkeit an. Örtlich kann der Einsatz

und des Bundesgerichtshofs, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 38.

entscheidungsersetzender informationstechnischer Systeme, vergleichbar der Konzentration bei Mahnverfahren, bei bestimmten Gerichten, auch länderübergreifend, zentral angesiedelt werden. Ausgesteuerte Verfahren können an das Gericht ausgesteuert werden, an dem das automatisierte Verfahren selbst erfolgt, an gut vernetzte Knotenpunkte oder an das Gericht, das zuständig wäre, wenn die Konzentrationsvorschrift nicht existierte. Aufgrund der Notwendigkeit eines Online-Verfahrens verbunden mit Barrierefreiheitsmaßnahmen kann auch Bedenken hinsichtlich des Zugangs zum Recht in der Fläche wirksam begegnet werden. Hinsichtlich der funktionellen Zuständigkeit im Aussteuerungsfalle bietet es sich an, die Fälle so zu verteilen, als wäre kein entscheidungsersetzendes System vorgesehen gewesen. Auch besteht insbesondere die Möglichkeit, die Fälle an mit sachnahen Aufgaben betraute Spruchkörper zuzuweisen.
Zuständigkeitsregelungen sind auch für die Aufsicht, Kontrolle und Korrektur des Systems festzulegen. Eine maßgebliche Handlungsoption besteht darin, das Vorhandensein einer interdisziplinären hohen Expertise für diese Tätigkeiten normativ festzulegen. Für die inhaltliche Ausgestaltung kommt eine Vielzahl an Möglichkeiten in Betracht. Auch die Konzentration der Ansiedelung der Aufsichts- und Kontrollinstitutionen ist in verschiedenen graduellen Ausgestaltungen möglich. Zudem sollten Potentiale aus den Fällen der Aussteuerung sowie eingelegter Rechtsmittel gegen entscheidungsersetzende informationstechnische Systeme genutzt werden.

b. Rote Linien für einen entscheidungsunterstützenden Einsatz sowie Handlungsleitlinien

aa. Rote Linien entscheidungsunterstützender Einsatz

Inwieweit für das „Ob“ und „Wie“ eines entscheidungs*unterstützenden* Einsatzes eine **Rechtsgrundlage** notwendig ist, hängt davon ab, welche Funktion das informationstechnische System erfüllt.[904] Normativ ist hierzu auf das Rechtsstaatsprinzip unter dem Gesichtspunkt des Vorbehalts des Gesetzes in Verbindung mit dem Wesentlichkeitsgrundsatz zu verweisen.[905] Zentral ist, inwieweit sich der entscheidungsunterstützende Einsatz auf die Verwirklichung der Grundrechte der Parteien auswirkt. Dafür ist die jeweilige Funktionalität des informationstechnischen Systems entscheidend. Übernimmt das System eine inhaltliche Funktion, kommt ihm eine stärkere **Grundrechtsrelevanz** zu. Jedenfalls für den Einsatz **inhaltsbezogener unterstützender Funktionen** bedarf es somit einer **gesonderten gesetzlichen Grundlage**.[906]

bb. Handlungsoptionen und -empfehlungen

Wie vorstehend dargestellt, lässt sich aus Art. 20 Abs. 3 GG ein Vorbehalt des Gesetzes für staatliches Handeln ableiten. In allen „wesentlichen“ Bereichen ist eine hinreichend bestimmte, durch den parlamentarischen Gesetzgeber geschaffene gesetzliche Grundlage erforderlich.[907] Der Gesetzgeber hat in grundlegenden

[904] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 152ff.
[905] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 152ff.
[906] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 152ff.
[907] *BVerfGE* 98, 218, 251, Rn. 136; Dreier/*Schulze-Fielitz*, 3. Aufl. 2015, GG Art. 20 Artikel 20 Rn. 105; Jarass/Pieroth/*Jarass*, 19. Aufl. 2026, GG Art. 20 Rn. 77.

normativen Bereichen die wesentlichen Entscheidungen selbst zu treffen (Wesentlichkeitsgrundsatz).[908] Da der Einsatz informationstechnischer Systeme zur Unterstützung der richterlichen Entscheidungsfindung durchaus erwünscht ist und auch in vielen Bereichen geradezu geboten erscheint, sollte eine Rechtsgrundlage für den Systemeinsatz (ggf. auch erst zum Zwecke der Erprobung) zeitnah geschaffen werden.

Eine entsprechende Rechtsgrundlage sollte explizit festhalten, dass der Einsatz informationstechnischer Systeme bei der richterlichen Entscheidungsfindung auch für inhaltsbezogene Funktionen, wie beispielsweise die Erstellung von Relationstabellen, das Zusammenfassen und das Generieren von Text zugelassen ist, sofern sichergestellt ist, dass die rechtlichen Anforderungen für einen solchen Einsatz eingehalten werden und hierbei insbesondere die berührten Interessen der Beteiligten hinreichend gewahrt werden sowie eine hinreichende Aufsicht und Kontrolle des eingesetzten Systems sichergestellt ist.

Dabei ist nicht erforderlich, dass die Rechtsgrundlage alle erdenklichen Einsatzszenarien ganz konkret umschreibt und die einsetzbaren Systeme bis in jedes Detail konkret individuell definiert. Aufgrund der rapiden Fortentwicklung der technischen Systeme und der Erschließung neuer Funktionalitäten wäre dies nicht zielführend, da mit Inkrafttreten der Regelung die Einsatzmöglichkeiten bereits derart fortentwickelt wären, dass das Recht mit Inkrafttreten bereits veraltet wäre. Es ist demgegenüber sinnvoll, abstrakt in einer Rechtsgrundlage die **konkreten Einsatzbereiche** (z. B. zivilgerichtliche Massenverfahren, etc.) und die in Betracht kommenden unterstützenden **Funktionalitäten exemplarisch** (Textgenerierung, Inhaltsextraktion etc.) mit abstrakt-generellen und technikoffenen Begrifflichkeiten zu definieren und somit einen rechtlichen Einsatzrahmen auch für künftig entwickelte Technologien zu schaffen. Hierbei sollten sowohl die inhaltsbezogenen Funktionen eine normative Regelung erfahren als auch die administrativen Funktionen, obgleich letztere streng genommen einer solchen nicht bedürfen. Durch die normative Verankerung der administrativen Funktionen erfolgt eine gesetzliche Klarstellung, die mit Blick auf die Rechtssicherheit und -klarheit zu begrüßen ist. Auch können bestimmte Funktionen je nach konkreter Ausgestaltung sowohl eine administrative Natur als auch inhaltsbezogene Natur aufweisen (Bsp.: Die Textaufbereitung kann in verschiedenen Formen erfolgen. Verlinkt ein System nur Rechtsprechung und Literatur, ist die Tätigkeit als administrativ einzustufen, markiert hingegen das System automatisch bestimmte relevante Passagen der Akte, ist dies als inhaltsbezogene Tätigkeit einzustufen.)[909]. Um zum einen die unterschiedliche Kategorisierung einer auf den ersten Blick gleichen Funktion den nutzenden Richterinnen und Richtern vor Augen zu führen sowie diese vor Rechtsverstößen zu bewahren, sollte im Einzelfall die konkrete Ausgestaltung der jeweiligen Funktion von einer administrativen Kategorisierung in eine inhaltsbezogene Kategorisierung umschlagen, ist eine derartige gesetzliche Regelung vorteilhaft. Auch kann zusätzlich die Ergänzung der gesetzlichen Regelung um eine Auffangklausel erweitert werden.

Die **Zulässigkeit des Einsatzes** von informationstechnischen Systemen für inhaltsbezogene Funktionen könnte **darüber hinaus an abstrakte Anforderungen geknüpft** sein (z. B. das Vorhandensein geeigneter Mechanismen zur Bewirkung effektiver menschlicher Aufsicht und Kontrolle; das Vorhandensein von Mechanismen zur Umsetzung datenschutzrechtlicher Vorgaben etc.), die jeweils bezogen auf konkret avisierte Funktionalitäten und Einsatzszenarien in einem gesonderten

[908] Dürig/Herzog/Scholz/*Grzeszick*, 111. EL Juni 2026, GG Art. 20 Rn. 105f.
[909] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 233f.

Anforderungsdokument bzw. **per Rechtsverordnung** einheitlich definiert werden könnten. Eine inhaltliche Orientierung für diese Anforderungen bietet beispielsweise bereits dieses Forschungsprojekt. Mit Blick auf den Wesentlichkeitsgrundsatz sind die entsprechenden elementaren Grundlagen hierbei in der parlamentsgesetzlichen Grundlage niederzulegen.[910]
Als Standort für eine entsprechende **bundeseinheitliche Regelung** böte sich das GVG an. Das GVG regelt die Grundlagen der Gerichtsbarkeit, regelt die Organisation sowie Aufgabenverteilung zwischen den Gerichten und die materiell-rechtlichen Grundlagen, welche die Rechtsprechung charakterisieren;[911] es entfaltet (unmittelbar oder mittelbar[912]) Wirkung in den einzelnen Gerichtsbarkeiten. Entscheidungsunterstützende informationstechnische Systeme gliedern sich thematisch an diese grundlegenden Regelungen der Rechtsprechungstätigkeit im GVG an. Auch vor dem Hintergrund, dass entscheidungsunterstützende Systeme nicht zwingend auf nur eine bestimmte Gerichtsbarkeit bzw. auf ein bestimmtes Rechtsgebiet ausgelegt sind, sondern durchaus auf mehrere Gerichtsbarkeiten bzw. Rechtsgebiete Anwendung finden können, wie etwa die diesem Forschungsprojekt entspringende MAKI-KI, die einen solchen generischen Ansatz verfolgt, ist eine Verortung der gesetzlichen Regelung im GVG zu befürworten.[913] Sofern konkrete Spezifika der einzelnen Gerichtsbarkeiten für ihren Aufgabenbereich konkretisierender bzw. abweichender Regelung bedürfen, können diese ebenfalls im GVG in entsprechender Gliederung oder in den jeweiligen Prozessordnungen geregelt werden. In der bundeseinheitlichen Regelung sollte zudem eine Verordnungsermächtigung zugunsten des Bundesjustizministeriums enthalten sein. Die konkrete Formulierung von Anforderungen per Rechtsverordnung ließe eine jederzeitige und schnelle Anpassung an weitergehende Bedürfnisse der Praxis bzw. geänderte Technologien zu. Eine Beschaffung der Software durch die Justizverwaltungen der Länder, wie sie bislang bereits Praxis ist, müsste aufgrund dieser Standardisierung der Anforderungen nicht überdacht werden. Mit Blick auf die Akzeptanz sowie die Bedürfnisabdeckung der Richterschaft sei jedoch auf die Vorteilhaftigkeit hingewiesen, die Richterschaft (über entsprechende Repräsentationsgremien) hierbei eng einzubinden.
Einer weitergehenden verfahrensrechtlichen Anpassung bedürfte es darüber hinaus nicht, da die Systeme von Richterinnen und Richtern in dem bereits jetzt bestehenden Zuständigkeits- und Organisationssystem genutzt würden. Der entscheidungsunterstützende Einsatz informationstechnischer Systeme bei der richterlichen Entscheidungsfindung beschränkt sich auf Hilfstätigkeiten im Vorfeld der Entscheidungsfindung, der Richter bzw. die Richterin trifft die Entscheidung letztendlich selbst.[914] Die derzeitigen Zuständigkeitsnormen der §§ 21e ff. GVG samt der auf ihnen beruhenden Geschäftsverteilungspläne regeln gerade diesen Einsatz menschlicher Richterinnen und Richter.[915] Bedacht werden sollte jedoch zusätzlich der Komplex der Aufsicht und Kontrolle der informationstechnischen Systeme beim entscheidungsunterstützenden Einsatz. Bei diesem Einsatzbereich kommen Systeme

[910] Dürig/Herzog/Scholz/*Grzeszick*, 111. EL Juni 2026, GG Art. 20 Rn. 105f.
[911] MüKoZPO/*Pabst*, 6. Aufl. 2022, GVG vor § 1 Rn. 1f.
[912] Siehe z. B. den Verweis in § 173 S. 1 VwGO.
[913] *Bundesministerium der Justiz und für Verbraucherschutz*, Massenverfahren mit KI schneller und effizienter bearbeiten, abrufbar unter https://www.bmjv.de/DE/themen/digitales/digitalisierung_justiz/digitalisierungsinitiative/laendervorhaben/_doc/artikel_vorhaben_13_MAKI.html (Abruf am 22.09.2026, 18:07 Uhr).
[914] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 221.
[915] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 221.

maschinellen Lernens und eine Vielzahl an Einsatzszenarien mit unterschiedlichen Graden an Komplexität, Funktionalität und Automatisierung in Betracht. Anhand der Bandbreite an Einsatzszenarien sowie der von den genannten Faktoren abhängigen Einsatzmatrix ergeben sich auch unterschiedlich geeignete Modelle bezüglich des Ausfüllens der Aufsichts- und Kontrollnotwendigkeit. So dürfte es ausreichen, den erkennenden Richter bzw. die erkennende Richterin mit der inhaltlichen Überwachung bei administrativen Funktionen zu betrauen (Bsp.: Ein agentischer KI-Verfahrensassistent benennt automatisiert eingehende Verfahrensinformationen [d. h. er weist Metadaten zu], ordnet die Information dem relevanten Verfahren zu und speichert sie in der elektronischen Akte an der relevanten Stelle abrufbar ab). Eine richterliche Einbindung in den Entwicklungsprozess wäre in diesem Fall nicht notwendig (gleichwohl vorteilhaft aufgrund der Expertise hinsichtlich der Verfahrensgestaltung). Während bei den administrativen Funktionen eine „Automatisierungskammer“ o. Ä. wie gerade aufgezeigt nicht notwendig ist, kommt derartiges jedoch bei (insbesondere stark) inhaltsbezogenen Unterstützungsfunktionen in Betracht (Bsp.: Das informationstechnische System gibt vollständige Entscheidungsvorschläge aus.).[916] Auch hierbei sind verschiedene Grade an Konzentration bzw. dezentraler Organisation denkbar (siehe hierzu bereits Abschnitt C.II.2.a.bb(1) sowie Abschnitt C.II.6.a.bb). Diese sind entsprechend auszugestalten und zu implementieren.
Abschließend sei noch angemerkt, dass es im entscheidungsunterstützenden Einsatzszenario keiner gesonderten datenschutzrechtlichen Rechtsgrundlage bedarf, da das derzeitige Recht entsprechende Rechtsgrundlagen für die Verarbeitung personenbezogener Daten bereits geschaffen hat, vgl. § 3 BDSG.[917]

cc. Zusammenfassung

Ob der entscheidungsunterstützende Einsatz einer Rechtsgrundlage bedarf, richtet sich nach dem Rechtsstaatsprinzip (Vorbehalt des Gesetzes i. V. m. dem Wesentlichkeitsgrundsatz) und hängt von der Grundrechtsrelevanz der jeweiligen Funktion ab. **Jedenfalls für inhaltsbezogene unterstützende Funktionen** bedarf es einer **gesonderten gesetzlichen Grundlage**.

Da der entscheidungsunterstützende Einsatz informationstechnischer Systeme erwünscht ist und oftmals geradezu geboten erscheint, **sollte** eine Rechtsgrundlage für den Systemeinsatz **zeitnah geschaffen werden**. Sie sollte explizit festhalten, dass auch der Einsatz inhaltsbezogener Funktionen zugelassen ist, sofern die notwendigen Anforderungen erfüllt sind. Die **Einsatzbereiche und Funktionalitäten** des informationstechnischen Systems sollten in der Rechtsgrundlage **abstrakt-generell und technikoffen umschrieben** werden und somit einen rechtlichen Einsatzrahmen auch für künftig entwickelte Technologien schaffen. Auch administrative Funktionen sollten, obwohl nicht zwingend nötig, klarstellenderweise normiert werden, u. a., da dieselbe Funktion je nach Ausgestaltung administrativer oder inhaltsbezogener Natur sein kann. Die Regelung kann durch eine Auffangklausel ergänzt werden. Die Zulässigkeit inhaltsbezogener Funktionen könnte an **abstrakte Anforderungen** geknüpft werden, die insbesondere **per Rechtsverordnung einheitlich definiert** werden könnten. Die **wesentlichen Anforderungen sind im Parlamentsgesetz**

[916] *Köhler*, Richterliche Verantwortung und Künstliche Intelligenz, 1. Aufl. 2026, S. 253 sieht im Falle von Automatisierungskammern eine rechtsgrundlagenbedürftige Delegation gegeben.
[917] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 101, 253f., 265ff., 331; siehe hierzu bereits zuvor im Abschnitt C.II.5.b.aa.

niederzulegen. Als Standort für eine bundeseinheitliche Regelung ist das **GVG** zu befürworten, da entscheidungsunterstützende Systeme in mehreren Gerichtsbarkeiten bzw. Rechtsgebieten Anwendung finden können. Spezifika einzelner Gerichtsbarkeiten können ebenfalls im GVG oder in den jeweiligen Prozessordnungen geregelt werden. Zudem sollte in der bundeseinheitlichen Regelung eine **Verordnungsermächtigung zugunsten des Bundesjustizministeriums** enthalten sein. Die Richterschaft sollte bei der Beschaffung der Systeme eng eingebunden sein. Einer weitergehenden verfahrensrechtlichen Anpassung bedürfte es darüber hinaus nicht, da die Systeme von Richterinnen und Richtern in dem bereits jetzt bestehenden Zuständigkeits- und Organisationssystem genutzt würden. Bedacht werden sollte jedoch der Komplex der Aufsicht und Kontrolle der informationstechnischen Systeme beim entscheidungsunterstützenden Einsatz. Die entsprechenden Aufsichts- und Kontrollnotwendigkeiten sind individuell anhand des konkreten Szenarios und der dreigliedrigen Funktionsmatrix zu bestimmen. Einer gesonderten datenschutzrechtlichen Rechtsgrundlage bedarf es für den entscheidungsunterstützenden Einsatz angesichts des § 3 BDSG nicht.

7. Verfügbarkeit von Rechtsbehelfen

Weitere rote Linien sind im ersten Arbeitspaket (Milestone 1) des Forschungsprojekts „Künstliche Intelligenz und richterliche Entscheidungsfindung“[918] mit Blick auf die Verfügbarkeit von Rechtsbehelfen identifiziert worden.

a. Übergreifende rote Linien für jeden Einsatz

Soweit es im Einzelfall zu einer Verletzung der genannten roten Linien kommt, ist ein angemessener, verfahrensmäßig abgesicherter Rechtsbehelf vorzusehen. Dies gilt v. a. für eine Verletzung von Grundrechten. Rechtsbehelfe können, müssen jedoch nicht gesondert gestaltet sein. Es ist mit den rechtlichen Vorgaben vereinbar, wenn eine entsprechende Rechtsverletzung über das herkömmliche Rechtsmittelverfahren geltend gemacht werden kann, d. h. im Falle eines entscheidungs*ersetzenden* Einsatzes in einem Rechtsmittel gegen die Entscheidung des informationstechnischen Systems oder im Falle eines entscheidungs*unterstützenden* Einsatzes gegen die Entscheidung, die unter Einsatz eines informationstechnischen Systems zustande gekommen ist bzw. über weitere existierende Rechtsmittel, die die Geltendmachung von Grundrechtsverstößen ermöglichen. Konkret zu verweisen ist neben den allgemeinen Gewährleistungen, wie dem Justizgewähranspruch gem. Art. 20 Abs. 3 GG i. V. m. Art. 2 Abs. 1 GG, auf Art. 13 EMRK sowie Art. 14 und 15 KI-Konvention.[919]

b. Handlungsoptionen und -empfehlungen

Bei einem entscheidungs*unterstützenden* Einsatz informationstechnischer Systeme trägt der menschliche Richter bzw. die menschliche Richterin die Letztverantwortung und es ergeben sich mit Blick auf Rechtsbehelfe sowie insbesondere **mit Blick auf Rechtsmittel keine Besonderheiten** im Vergleich zum herkömmlichen Zivilverfahren.[920] Dort lassen sich Rechts- und v. a. Grundrechtsverletzungen vorbringen und in einem angemessenen Verfahren behandeln. Gleichwohl werden auch in diesem Zusammenhang Vorschläge zur Ergänzung des Rechtsbehelfskanons gemacht. Vorgeschlagen wird z. B. die Schaffung eines Rechtsbehelfs sui generis zur Korrektur KI-bedingter Fehler, angelehnt an die Gehörsrüge nach § 321a ZPO, gleichwohl um die Subsidiarität gegenüber anderen Rechtsbehelfen enthoben.[921] Ein weiterer Vorschlag ist die Zuleitung des fertigen Entscheidungsentwurfs an die Parteien zur Prüfung auf KI-bedingte Fehler, wobei dies zur Vermeidung des Missbrauchs als vorgelagerte Rechtsmittelinstanz mit einem breiten Zurückweisungsrecht kombiniert werden soll.[922] Beide Vorschläge sollen der Korrektur KI-bedingter Fehler sowie der Sicherstellung richterlicher Kontrolle dienen; bereits die Einräumung dieser Rechtsbehelfe soll die adäquate richterliche Arbeit präventiv positiv beeinflussen.[923] Inwieweit jedenfalls jedoch zumindest die Zuleitung des fertigen Entscheidungsentwurfs selbst durch den Richter bzw. die Richterin an die Parteien vor Entscheidungserlass, nachdem der Richter bzw. die Richterin in den Entwurf bereits

---

[918] Siehe hierzu ausführlich *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025.
[919] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 251f., 260.
[920] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 90f., 108ff., 119f., 154f., 168, 177ff., 240f., 251f., 255f., 265ff., 344f., 358f., vgl. auch S. 236.
[921] *Gsell/Pfündl*, RDi 2026, 457, 465.
[922] *Gsell/Pfündl*, RDi 2026, 457, 465.
[923] *Gsell/Pfündl*, RDi 2026, 457, 464f.

seine/ihre Arbeitskraft investiert und diesen für präsentabel haltend nach außen gegeben hat, praxiswirksam tatsächlich zu einer materiellen Änderung des Entscheidungsinhalts im Rahmen der endgültigen Entscheidung führt, erscheint offen. **Handlungsnotwendigkeiten** mit Blick auf den entscheidungsunterstützenden Einsatz **ergeben sich indes nicht**.

Mit Blick auf einen entscheidungser*setzenden* Einsatz informationstechnischer Systeme gilt dies, soweit die Einbettung in das **bestehende Rechtsmittelsystem** erfolgt, ebenfalls, denn Rechts- und Grundrechtsverletzungen könnten dann im Rahmen des Rechtsmittelverfahrens vorgebracht werden. Dies zeigt sich etwa an den Berufungs- und Revisionsgründen, vgl. §§ 513 Abs. 1 (i. V. m.) § 546 ZPO und § 529 ZPO, die eine Geltendmachung von Rechtsverletzungen und die Berücksichtigung fehlerhafter Tatsachenfeststellungen in hinreichend umfangreichem Ausmaß gestatten. Liegt eine Rechtsverletzung durch ein informationstechnisches System vor, kann dies somit beispielsweise im Rahmen der Berufung gerügt werden, so dass eine konkrete Anpassung der Berufungsgründe auch im Falle der Einführung einer weiteren automatisierten Instanz nicht erforderlich erscheint.[924]

Ergänzend sind **weitere Gestaltungsoptionen** denkbar. Sollte die Ausgestaltung des entscheidungsersetzenden informationstechnischen Systems ähnlich dem Mahnverfahren gem. §§ 688 ff. ZPO verfolgt werden, könnte darin eine **auf Antrag zu gewährende Aussteuerungsmöglichkeit** eingeräumt werden, die den **Rechtsstreit einem menschlichen Richter bzw. einer menschlichen Richterin zuweist**.[925] Infolge der Einlegung eines Äquivalents zum Widerspruch bzw. Einspruch könnte sodann ähnlich der §§ 696 f., 700 ZPO ein streitiges gerichtliches Verfahren vor einem menschlichen Richter bzw. einer menschlichen Richterin durchgeführt werden.[926] Diese Rechtsbehelfe führen – unter Einhaltung der entsprechenden (weiteren) Voraussetzungen – dazu, dass durch Einleitung des streitigen Verfahrens (§ 696 Abs. 1 ZPO bzw. § 700 Abs. 3 ZPO) sich eine menschliche Richterin bzw. ein menschlicher Richter des Verfahrens annimmt.[927] Auf diese Weise werden zwar streng genommen keine konkreten Rechtsverletzungen des automatisierten Verfahrens gerügt, im sich anschließenden streitigen Verfahren vor einem menschlichen Richter bzw. einer menschlichen Richterin könnten potentielle Rechtsverletzungen allerdings dann durch entsprechendes Vorbringen der Parteien vermieden werden. Zu denken ist etwa an besondere Verteidigungsmittel, die im automatisierten Verfahren nicht erhoben werden können. Eine entsprechende Option der Parteien zur Aussteuerung ist daher durchaus sinnvoll.

Sollte die Ausgestaltung des entscheidungsersetzenden informationstechnischen Systems als – dem Rechtsprechungsbegriff unterfallende – vollwertige gerichtliche Instanz vorgenommen werden, gelten freilich strengere Anforderungen.[928] Sollte

---

[924] Zur Berufung als Fehlerkontrolle BeckOK ZPO/*Wulf/Gaier*, 61. Ed. 1.7.2026, ZPO § 513 Rn. 1ff.; Saenger/*Wöstmann*, 10. Aufl. 2023, ZPO § 513 Rn. 1.

[925] Vgl. *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 116f.

[926] Vgl. *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 116f.

[927] Thomas/Putzo/*Nordmeier* ZPO, 47. Aufl. 2026, § 694 Rn. 1; MüKoZPO/*Schüler*, 7. Aufl. 2025, ZPO § 694 Rn. 1f., 20 (Der Autor spricht a. a. O. sogar von einer „Rechtsmittelfrist“.), § 696 Rn. 17, § 697 Rn. 14; BeckOK ZPO/*Dörndorfer*, 61. Ed. 1.7.2026, ZPO § 700 Rn. 4ff., 7ff.; BeckOK ZPO/*Toussaint*, 61. Ed. 1.7.2026, ZPO § 338 vor Rn. 1, Rn. 1. Vgl. auch § 20 Abs. 1 Nr. 1 letzter Hs. RPflG. Vgl. auch *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 174ff. Eine solche Ausgestaltung als freiwillige Option befürwortend, sofern kein Verbraucher gegen einen Unternehmer klagt *Quack*, Automation zivilrechtlicher Entscheidungsfindung, 1. Aufl. 2025, S. 242ff.

[928] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 174f., 157ff.; ein automatisiertes Vorentscheidungsverfahren (als Art erweitertes Mahnverfahren verstehend), aber auch in der Art einer gerichtlichen Vorentscheidung wird als im Widerspruch zum zentralen Angebot

hierbei eine Verletzung der genannten roten Linien eintreten, ist auch in diesem Fall ein angemessener verfahrensmäßig abgesicherter Rechtsbehelf vorzusehen. Bei einer Einbindung in das System der **berufungsrechtlichen Regelungen**, insbesondere des § 511 ZPO, wäre dies unproblematisch der Fall. Modifikationen bedürfte das geltende Recht allerdings jedenfalls bei der Einführung einer echten „0.-Instanz“, da diese Instanzstufe bislang nicht vorgesehen ist:
Die Berufung ist gem. § 511 Abs. 1 ZPO statthaft gegen die im ersten Rechtszug erlassenen Endurteile der Amts- und Landgerichte.[929] Dem Grunde nach dürfte in dieser Hinsicht kein Anpassungsbedarf bestehen. Die in einer automatisierten, vollwertigen gerichtlichen („0.-“ oder ersten) Instanz ergangenen Entscheidungen mittels entscheidungsersetzender informationstechnischer Systeme wären berufungsfähige Endurteile i. S. v. § 300 ZPO.[930] Sollte die automatisierte Instanz bei den Amts- bzw. Landgerichten angesiedelt sein, wäre auch insoweit der Tatbestand des § 511 Abs. 1 ZPO erfüllt. Bei einer anderweitig angesiedelten – etwa einer in Bundeshand liegenden – Zuständigkeit wäre eine entsprechende Modifikation zur Einbeziehung in die Statthaftigkeit der Berufung vorzunehmen.
Für bürgerlich-rechtliche Streitigkeiten, die vor den Amtsgerichten verhandelt wurden (vgl. §§ 23 ff. GVG), sind die Zivilkammern der Landgerichte grundsätzlich gem. § 72 Abs. 1 S. 1 GVG die funktionell zuständigen Berufungsgerichte.[931] Die funktionelle Zuständigkeit für die Berufung gegen Entscheidungen der Landgerichte in Zivilsachen liegt gem. § 119 Abs. 1 Nr. 2 GVG bei den Oberlandesgerichten. Sofern eine echte „0.-Instanz“ eingeführt werden sollte, müsste angesichts dieser Zuständigkeitsregelung eine weitere Berufungszuständigkeit auf Ebene der ersten Instanz (für Berufungen gegen Entscheidungen der „0. Instanz“) eingezogen werden. Dies erscheint möglich und sinnvoll, so dass beispielsweise bei den Amtsgerichten die Berufungszuständigkeit für Entscheidungen über Berufungen gegen die „0.-Instanz“ liegen könnte. Soweit die erstinstanzlichen Urteile – z. B. für sämtliche Verfahren einer bestimmten Verfahrenskategorie[932] – entscheidungsersetzend von den informationstechnischen Systemen vorgenommen werden, müssen diese ja letztlich nicht mehr von (den) menschlichen Richterinnen und Richtern an den Amtsgerichten bearbeitet werden, die zuvor für die entsprechenden Bereiche zuständig waren. Um diese Expertise gleichwohl zu nutzen – sowie dem Leitbild einer „0.-Instanz“ zu folgen –, könnte diesen Gerichten die Zuständigkeit für die berufungsrechtliche Überprüfung der Entscheidungen zugewiesen werden, die seitens der informationstechnischen Systeme entscheidungsersetzend erfolgt sind. Die Amtsgerichte würden damit nunmehr neuerdings die Aufgaben eines Berufungsgerichts wahrnehmen. Damit müssten die Gerichte und ihre Richter sich die entsprechenden Kompetenzen – soweit nicht ohnehin schon vorhanden – aneignen. Lehrgänge sowie juristische Ausbildungsliteratur dürften hierbei eine grundlegende Unterstützung bieten.[933]

---

des Staates und der Justiz zur Konfliktlösung durch (menschliche) Richterinnen und Richter nach rechtsstaatlichen Maßstäben abgelehnt durch die *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, Münchener Thesen zum Zivilprozess der Zukunft, abrufbar unter https://www.olg-duesseldorf.nrw.de/infos/Zivilprozess-der-Zukunft/Muenchener-Thesen-zum-Zivilprozess-der-Zukunft.pdf (Abruf am 26.02.2026, 18:27 Uhr), S. 6, These A.5.

[929] Siehe auch MüKoZPO/*Rimmelspacher*, 7. Aufl. 2025, ZPO § 511 Rn. 19.

[930] Zur Berufungsfähigkeit der Endurteile MüKoZPO/*Rimmelspacher*, 7. Aufl. 2025, ZPO § 511 Rn. 6ff.

[931] *Adolphsen*, Zivilprozessrecht, 9. Aufl. 2025, H. § 30 Rn. 6f.

[932] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 219.

[933] Vgl. z. B. *Oberheim*, Zivilprozessrecht für Referendare, 15. Aufl. 2023, Rn. 1484ff., 1497ff., 1513ff.; *Rosenberg/Schwab/Gottwald*, Zivilprozessrecht, 18. Aufl. 2018, §§ 134ff., 137ff.

Ausgehend davon, an welchem Eingangsgericht mit der automatisierten Instanz der Rechtsstreit in den Instanzenzug eintritt, könnte das gleiche Gericht – indes menschlich besetzt – auch zugleich Berufungsgericht sein oder die Berufungszuständigkeit ist wie bislang auch bei dem Gericht der nächsthöheren Instanz angesiedelt. Die Revisionszuständigkeit kann wie bislang stets beim Bundesgerichtshof verortet werden.[934] Es wäre jedoch auch möglich, die Revisionszuständigkeit gegen die erstinstanzlich automatisiert gefassten Entscheidungen den Oberlandesgerichten zuzuordnen. Dies wäre für die Zivilgerichtsbarkeit ein Novum, ist jedoch bereits aus dem Bereich des Strafprozesses bekannt.[935] Auch sollte hier u. a., entsprechend der strafprozessualen Handhabe, nach § 121 Abs. 2 GVG eine Vorlagepflicht an den Bundesgerichtshof vorgesehen werden, wenn ein Oberlandesgericht von der Rechtsprechung eines anderen oder des Bundesgerichtshofs abweichen will.[936] Eine ähnliche Konstellation ist auch schon im Zusammenhang der Revisionszulassungsgründe gem. § 543 Abs. 2 S. 1 Nr. 2 Alt. 2 ZPO bekannt.[937] Es bestehen freilich verschiedene Kategorisierungsmöglichkeiten dafür, welche Rechtsstreitigkeiten einer solchen oberlandesgerichtlichen Revision zugewiesen werden. Eine Möglichkeit wäre beispielsweise, die erstinstanzlich den Amtsgerichten zugewiesenen Streitigkeiten revisionsrechtlich den Oberlandesgerichten zuzuordnen.[938] Auch streitwertabhängige oder streitwertunabhängige Zuweisungen mit der Festlegung einer grundsätzlichen Allzuständigkeit für Revisionen, vergleichbar der erstinstanzlichen Zuständigkeitsbestimmung wären denkbar.[939] Eine Sprungrevision kann im Fall der entscheidungsersetzenden Eingangsinstanz als eher kritisch beachtet werden, da auf den dann für die Sprungrevision feststehenden Sachverhalt noch kein menschlicher Richter bzw. keine menschliche Richterin Einfluss hatte.[940]

---

[934] *Adolphsen*, Zivilprozessrecht, 9. Aufl. 2025, § 31 Rn. 3.

[935] *Adolphsen*, Zivilprozessrecht, 9. Aufl. 2025, § 31 Rn. 3; *Ostendorf/Brüning*, Strafprozessrecht, 5. Aufl. 2024, § 17 Rn. 39; die Erwägung der Zuständigkeit von Oberlandesgerichten als Revisionsinstanz ist nicht vollständig fremd, vgl. BT-Drucks. 7/2459. In diesem Gesetzentwurf der CDU/CSU vom 07.08.1974, der sich durch den Ablauf der Wahlperiode erledigte (vgl. *Deutscher Bundestag*, Gesetz zur Einführung der Zulassungsrevision gegen Berufungsurteile der Landgerichte in Zivilsachen, abrufbar unter https://dip.bundestag.de/vorgang/gesetz-zur-einf%C3%BChrung-der-zulassungsrevision-gegen-berufungsurteile-der-landgerichte-in/305207 [Abruf am 25.02.2026, 17:47 Uhr]), wurde erwogen, den Rechtsstreitigkeiten, die in der ersten Instanz von den Amtsgerichten und in der Berufungsinstanz von den Landgerichten entschieden werden, die Revision vor den Oberlandesgerichten zu eröffnen. Die Ausgangslage war hier indes, dass gegen derartige Entscheidungen keine Revision möglich war, siehe etwa Stein/*Roth*, 24. Aufl. 2024, ZPO § 1 Rn. 59 m. w. N.

[936] *Ostendorf/Brüning*, Strafprozessrecht, 5. Aufl. 2024, § 17 Rn. 39; in dem Gesetzesentwurf der CDU/CSU vom 07.08.1974, BT-Drucks. 7/2459 war eine solche Möglichkeit ebenfalls vorgesehen.

[937] Vgl. BeckOK ZPO/*Kessal-Wulf/Rensen*, 61. Ed. 1.7.2026, ZPO § 543 Rn. 26.

[938] Diesen Ansatz auch primär verfolgend BT-Drucks. 7/2459; dies ist cum grano salis auch vergleichbar der strafprozessualen Ausgestaltung, bei der Oberlandesgerichte für Revisionen gegen erstinstanzliche amtsgerichtliche Urteile sowie landgerichtliche Berufungsurteile zuständig sind, vgl. *Kindhäuser/Schumann*, Strafprozessrecht, 8. Aufl. 2025, § 31 Rn. 5.

[939] Vgl. MüKoZPO/*Pabst*, 6. Aufl. 2022, GVG § 23 Rn. 1; Musielak/Voit/*Wittschier*, 23. Aufl. 2026, GVG § 23 Rn. 1; eine Streitwertrevision war in der Bundesrepublik Deutschland bis zum Jahr 2001 gegeben, sie wurde jedoch mit dem Gesetz zur Reform des Zivilprozesses vom 27.07.2001 – BGBl. I 2001, Nr. 40 vom 02.08.2001, S. 1887ff. abgeschafft, vgl. BT-Drucks. 14/4722, S. 1: „Die Reform enthält folgende Schwerpunkte: (…) Abschaffung der Streitwertrevision, (…)“, S. 103: „Das bisherige Mischsystem von Zulassungs- und Streitwertrevision wird durch eine generelle Zulassungsrevision mit Nichtzulassungsbeschwerde ersetzt und vereinheitlicht den Zugang zur Revisionsinstanz.“.

[940] *Quack*, Automation zivilrechtlicher Entscheidungsfindung, 1. Aufl. 2025, S. 242. Hiergegen kann jedoch vorgebracht werden, dass die Sprungrevision de lege lata der Einwilligung des Gegners gem. § 566 Abs. 1 S. 1 ZPO bedarf. Sind sich beide Parteien über den durch das informationstechnische

Mit dem Eingliedern der „0.-Instanz“ in den Instanzenzug würde der bisherige dreigliedrige Instanzenzug um eine vorgeschaltete vierte Instanz erweitert werden.[941] Damit wäre für den Rechtsschutzsuchenden kein Instanzverlust, sondern lediglich eine Erweiterung mit Blick auf den Instanzenzug gegeben, was sich grundsätzlich akzeptanzfördernd auswirken dürfte. Ob dies indes (stets) notwendig ist, kann jedoch in Frage gestellt werden. Für die (hinreichend einfach gelagerten[942]) Fälle, in denen derzeit ein entscheidungsersetzender Einsatz lediglich in Betracht kommt, erscheint angesichts der Komplexität und der Streitgegenstände ein viergliedriger Instanzenzug eher unangemessen. Es ist daher durchaus eine Möglichkeit, in diesen unkomplexen Fällen den dreigliedrigen Instanzenzug – jedoch mit dem entscheidungsersetzenden informationstechnischen System als erste Instanz – beizubehalten.[943] Mit Blick auf die Zukunft und den technischen Fortschritt könnte bei entsprechender Erhöhung der Leistungsfähigkeit der Systeme sich dieser Befund jedoch ändern, indem eine – ggf. sogar überwiegende – Mehrzahl an Verfahren vollautomatisiert von entscheidungsersetzenden informationstechnischen Systemen bearbeitet wird. Dies ist jedoch Zukunftsmusik und liegt derzeit noch in weiter Ferne.

Entsprechend der (genauen) Ausgestaltung des statthaften Rechtsbehelfs gegen Entscheidungen, die durch informationstechnische Systeme getroffen wurden, sind auch die **Rechtsbehelfsbelehrungen** gem. § 232 ZPO anzupassen.[944]

Sowohl die Berufungseinlegung gem. § 519 Abs. 1 ZPO als auch die Berufungsbegründung gem. § 520 Abs. 3 S. 1 ZPO sind bei dem Berufungsgericht, dem sog. iudex ad quem, einzureichen.[945] Eine Ausgestaltung dergestalt, dass das Gericht, bei dem die Entscheidung in vollständig automatisierter Weise getroffen wurde, zugleich die Rechtsmittelinstanz ist, hätte den Vorteil, dass insoweit Probleme im Rahmen und in der Folge der Einreichung verringert werden können.[946]

Bezüglich des Inhalts der Berufungsbegründung gem. § 520 Abs. 3 S. 2 ZPO geraten insbesondere Nr. 2 und Nr. 3 ins Augenmerk: § 520 Abs. 3 S. 2 Nr. 2 ZPO verlangt die Bezeichnung der Umstände, aus denen sich die Rechtsverletzung und deren Erheblichkeit für die angefochtene Entscheidung ergeben. Die Rechtsverletzung, mithin Interpretationsfehler oder Subsumtionsfehler, die zur inhaltlichen Unrichtigkeit des angegriffenen Urteils geführt haben, muss angegeben werden.[947] Hierbei bedarf es einer auf den Streitfall zugeschnittenen Darstellung. Eine Rüge der Auffassung des Erstgerichts mit formularmäßigen Sätzen oder allgemeinen Redewendungen bzw. nur der Verweis auf das Vorbringen der Vorinstanz genügen nicht.[948] Diese Anforderungen gilt es freilich (neben den weiteren) auch im Rahmen der Berufungsschriften gegen automatisiert erlassene Urteile zu wahren. Pauschale, auf den Einsatz informationstechnischer Systeme bezogene Floskeln, die z. B. darauf abstellen, das informationstechnische System hätte im vorliegenden Fall nicht richtig funktioniert, würden demnach nicht ausreichen. Es ist weiterhin eine hinreichend auf den Einzelfall bezogene Darstellung notwendig.

System festgestellten Sachverhalt einig (nur dann wird der Gegner in die Sprungrevision einwilligen), bedürfte es einer tatsächlichen Überprüfung gerade nicht.

941 *Quack*, Automation zivilrechtlicher Entscheidungsfindung, 1. Aufl. 2025, S. 241.

942 Vgl. *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 87ff., 106ff., 115ff., 143ff., 170ff., 229ff., 240ff., 255f., (258ff.,) 265ff., 334f., 338f., 344, 356ff.

943 Eher ablehnend *Quack*, Automation zivilrechtlicher Entscheidungsfindung, 1. Aufl. 2025, S. 241f.

944 Vgl. zum Inhalt einer Rechtsbehelfsbelehrung Stein/*Loyal*, 24. Aufl. 2024, ZPO § 232 Rn. 12.

945 *Adolphsen*, Zivilprozessrecht, 9. Aufl. 2025, § 30 Rn. 12, 21ff.

946 Zu derartigen Problematiken bspw. *Adolphsen*, Zivilprozessrecht, 9. Aufl. 2025, § 30 Rn. 12.

947 Thomas/Putzo/*Seiler*, ZPO, 47. Aufl. 2026, § 520 Rn. 22; Thomas/Putzo/*Seiler*, ZPO, 47. Aufl. 2026, § 546 Rn. 1ff.

948 Thomas/Putzo/*Seiler*, ZPO, 47. Aufl. 2026, § 520 Rn. 22.

§ 520 Abs. 3 S. 2 Nr. 3 ZPO bedarf insoweit der Bezeichnung konkreter Anhaltspunkte, die Zweifel an der Richtigkeit oder Vollständigkeit der Tatsachenfeststellungen im angefochtenen Urteil begründen und deshalb eine erneute Feststellung gebieten. Hiermit soll insbesondere bloß formelhaften Berufungsbegründungen entgegengewirkt werden.[949] Es bedarf konkreter Angaben, die Darlegungen müssen auf den konkreten Fall zugeschnitten sein.[950] In Betracht kommen insbesondere Verfahrensfehler, namentlich Verletzungen des Rechts auf rechtliches Gehör gem. Art. 103 Abs. 1 GG.[951] Hierbei sind eine Vielzahl von Fehlern denkbar, die insbesondere auch aufgrund von (spezifischen Fehlern der) eingesetzten informationstechnischen Systemen bei der richterlichen Entscheidungsfindung Relevanz haben können, wie beispielhaft das nicht (vollständige) Berücksichtigen von Tatsachen- oder Beweisvorbringen.[952] Die Anforderungen der Berufungsbegründung sind auch im Kontext des Einsatzes informationstechnischer Systeme bei der richterlichen Entscheidungsfindung zu wahren.

Anpassungsbedarf gäbe es im Recht der Berufung allerdings in einem konkreten Szenario: Hinsichtlich der Statthaftigkeit der Berufung wird zwischen der Wertberufung nach § 511 Abs. 2 Nr. 1 ZPO und der Zulassungsberufung nach § 511 Abs. 2 Nr. 2 ZPO unterschieden. Als hinreichend einfach gelagerter Komplex für den Einsatz entscheidungsersetzender informationstechnischer Systeme wurden bereits Ausgleichsansprüche nach der Fluggastrechte-VO identifiziert.[953] Mit Blick auf deren (singulär betrachtet) maximale Höhe von 600 Euro nach Art. 7 Abs. 1 S. 1 lit. c) Fluggastrechte-VO, werden dies betreffende Verfahren (regelmäßig) nicht die Beschwerdegegenstands-Wertgrenze von 1000 Euro übersteigen.[954] In solchen Fällen bedürfte es daher der Zulassungsberufung gem. § 511 Abs. 2 Nr. 2, Abs. 4 ZPO. Die Rechtsmittelzulassung ist jedoch aufgrund ihrer Unbestimmtheit und uneinheitlichen Handhabe Kritik ausgesetzt.[955] Das gesellschaftliche Grundvertrauen in den Rechtsstaat und in die Fähigkeit der Gerichte, (hochwertige) Lösungen bereitzustellen, sind Fundamente der unabhängigen Judikative eines demokratischen Rechtsstaats.[956] Aus diesem Grund erscheint es vorteilhaft, dieses Grundvertrauen zu erhalten und zu unterstützen, indem **Berufungen gegen Entscheidungen, die durch entscheidungsersetzende informationstechnische Systeme getroffen wurden**, nicht dem Weg über eine problembehaftete Rechtsmittelzulassung auszusetzen, sondern die Berufung (auch) in derartigen Fällen **stets zuzulassen**.

---

[949] *Adolphsen*, Zivilprozessrecht, 9. Aufl. 2025, § 30 Rn. 25.
[950] Thomas/Putzo/*Seiler*, ZPO, 47. Aufl. 2026, § 520 Rn. 20, 23.
[951] BeckOK ZPO/*Gaier*, 61. Ed. 1.7.2026, ZPO § 520 Rn. 35; MüKoZPO/*Rimmelspacher*, 7. Aufl. 2025, ZPO § 520 Rn. 51.
[952] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 229ff.; MüKoZPO/*Rimmelspacher*, 7. Aufl. 2025, ZPO § 520 Rn. 53ff.
[953] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, vgl. nur S. 10ff., 145.
[954] Zum Begriff des Beschwerdegegenstands siehe BeckOK ZPO/*Wulf*/*Gaier*, 61. Ed. 1.7.2026, ZPO § 511 Rn. 35.
[955] Vgl. nur *Greger*, Stellungnahme zum Entwurf eines Gesetzes zur Änderung der Gesetzes [sic], betreffend die Einführung der Zivilprozessordnung (BT-Drucks. 19/1686) vom 08.05.2018, abrufbar unter https://www.bundestag.de/resource/blob/554494/c352b8830ab4c3abe77e3c3d2b32aa94/greger-data.pdf (Abruf am 17.03.2026, 13:58 Uhr), S. 33.
[956] Die *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, Münchener Thesen zum Zivilprozess der Zukunft, abrufbar unter https://www.olg-duesseldorf.nrw.de/infos/Zivilprozess-der-Zukunft/Muenchener-Thesen-zum-Zivilprozess-der-Zukunft.pdf (Abruf am 26.02.2026, 18:27 Uhr), S. 2.

Mit Blick auf die umfangreich zugelassenen Gründe für eine Berufung besteht jedoch **kein grundlegender Handlungsbedarf**. Die **derzeitigen Berufungsgründe** erscheinen auch für die Überprüfung von Entscheidungen, die durch entscheidungsersetzende informationstechnische Systeme getroffen wurden, **tauglich**.
Als **möglicher weiterer, gleichwohl nicht zwingender Rechtsbehelf** könnte auch ein im System bereits vorgesehenes sog. **Flagging-Tool** (i. S. e. Schaltfläche innerhalb des Justizportals) angedacht werden, mit dem Grundrechts- bzw. Gewährleistungsverletzungen seitens eines entscheidungsersetzenden informationstechnischen Systems geltend gemacht werden könnten.[957] Der Zeitpunkt der Geltendmachung kann hierbei zu verschiedenen Verfahrensstadien eröffnet werden, insbesondere noch im Rahmen des laufenden Verfahrens oder nach Abschluss des Verfahrens (z. B. nach Ergehen des Urteils). Ein Flagging würde die Überprüfung des Systems bzw. der Systemausgabe durch einen menschlichen Richter zur Folge haben.[958] Eine Ausgestaltung dergestalt, dass ein Flagging vor Rechtskraft der Entscheidung eröffnet wird, könnte vorteilhaft sein, da sie eine schnelle Korrektur bzw. Validierung einer Entscheidung ermöglicht, ohne dass der Instanzenzug beschritten werden müsste, wobei dieser unabhängig von dieser Rechtsbehelfsmöglichkeit (insoweit unverändert) bestehen und möglich sein sollte.[959]

c. Zusammenfassung

Kommt es zu einer Verletzung der roten Linien, insbesondere von Grundrechten, ist ein angemessener, verfahrensmäßig abgesicherter Rechtsbehelf vorzusehen. Es muss hierfür nicht eigens ein neuer Rechtsbehelf geschaffen werden. Ausreichend ist, wenn der Verstoß über die bestehenden Rechtsmittel geltend gemacht werden kann.

Bei einem **entscheidungsunterstützenden Einsatz** trägt die Richterin bzw. der Richter die Letztverantwortung. Daher ergeben sich **weder Besonderheiten noch Handlungsnotwendigkeiten** (auch wenn in der Literatur Vorschläge zur Ergänzung des Rechtsbehelfskanons gemacht werden) für den entscheidungsunterstützenden Einsatz **mit Blick auf Rechtsmittel**. Gleiches gilt grundsätzlich für den entscheidungsersetzenden Einsatz informationstechnischer Systeme, soweit dieser in das bereits bestehende Rechtsmittelsystem eingebettet wird. Die Berufungs- und Revisionsgründe, vgl. §§ 513 Abs. 1 (i. V. m.) § 546 ZPO und § 529 ZPO, gestatten bereits die Geltendmachung von Rechtsverletzungen und die Berücksichtigung fehlerhafter Tatsachenfeststellungen in hinreichend umfangreichem Ausmaß. Bei einer Gestaltung ähnlich dem Mahnverfahren kann die Gewährung einer Aussteuerungsmöglichkeit auf Antrag, die den Rechtsstreit einem menschlichen Richter bzw. einer menschlichen Richterin zuweist, verfolgt werden. Ähnlich zum Widerspruch bzw. Einspruch könnte das Verfahren so in ein streitiges gerichtliches Verfahren vor einem menschlichen Richter bzw. einer menschlichen Richterin übergeleitet werden. Bei Ausgestaltung als vollwertige gerichtliche Instanz gelten strengere Anforderungen. Die Einbindung in das System der berufungsrechtlichen Regelungen ist unproblematisch, bedarf jedoch ggf. Modifikationen. Die Entscheidungen des informationstechnischen Systems wären berufungsfähige Endurteile. Liegt die Zuständigkeit nicht bei den Amts- oder Landgerichten, wäre die Regelung der Statthaftigkeit der Berufung anzupassen. Bei Einführung einer „0.-

[957] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 264.
[958] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 264.
[959] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 264.

Instanz“ bedarf es einer neuen Berufungszuständigkeit auf der Ebene der ersten Instanz, welche zum Beispiel bei den Amtsgerichten angesiedelt werden könnte. Die Revisionszuständigkeit kann beim Bundesgerichtshof verbleiben oder mit einer Vorlagepflicht entsprechend § 121 Abs. 2 GVG den Oberlandesgerichten zugewiesen werden. Eine Sprungrevision kann kritisch betrachtet werden, da auf den dann für die Sprungrevision feststehenden Sachverhalt noch kein menschlicher Richter bzw. keine menschliche Richterin Einfluss hatte. Für die derzeit in Betracht kommenden hinreichend einfach gelagerten Fälle erscheint ein viergliedriger Instanzenzug eher unangemessen. Insoweit kann der dreigliedrige Instanzenzug beibehalten werden. Die geltenden Berufungsgründe weisen keinen grundlegenden Handlungsbedarf auf. Anzupassen sind u. U. die Rechtsbehelfsbelehrungen. Anpassungsbedarf besteht indes bei der Statthaftigkeit: Verfahren über geringe Geldsummen, wie im Falle von Ausgleichsansprüchen nach der Fluggastrechte-VO, werden die Beschwerdegegenstands-Wertgrenze von 1000 Euro regelmäßig nicht überschreiten. Zur Wahrung insbesondere des Grundvertrauens in den Rechtsstaat sollten Berufungen gegen Entscheidungen, die durch entscheidungsersetzende informationstechnische Systeme getroffen wurden, stets zugelassen werden, statt sie dem Weg über eine problembehaftete Rechtsmittelzulassung auszusetzen. **Optional** könnte **als weiterer Rechtsbehelf ein Flagging-Tool im Justizportal** vorgesehen werden. Über dieses könnten Grundrechts- bzw. Gewährleistungsverletzungen seitens des Systems geltend gemacht werden und eine Überprüfung durch einen menschlichen Richter bzw. eine menschliche Richterin ausgelöst werden. Vorteilhaft wäre die Eröffnung des Rechtsbehelfs bereits vor der Rechtskraft. Der Instanzenzug bliebe davon unberührt.

8. Barrierefreiheit des Rechtsschutzes

Anforderungen ergeben sich auch mit Blick auf den Zugang zum Recht, nachgerade die Barrierefreiheit des Rechtsschutzes.

a. Übergreifende rote Linien für jeden Einsatz

In unserem ersten Arbeitspaket (Milestone 1) des Forschungsprojekts „Künstliche Intelligenz und richterliche Entscheidungsfindung“[960] ist eingehend herausgearbeitet worden, dass alle Rechtsschutzsuchenden angemessenen **Zugang zu Gericht** erhalten müssen. Findet der Zugang zu einem Verfahren im entscheidungsersetzenden Einsatz vornehmlich digital über eine Online-Plattform statt, sind für Personen ohne die notwendigen digitalen Kompetenzen oder Ressourcen (Internetzugang) bei den Gerichten niedrigschwellige Unterstützungsmöglichkeiten einzurichten.[961]
Dies folgt letztlich aus dem Justizgewährungsanspruch gemäß Art. 2 Abs. 1, 20 Abs. 3 GG[962], dem Anspruch auf ein faires Verfahren[963], Art. 6 Abs. 1 EMRK[964] und Art. 47 Abs. 1, Abs. 2 GRCh[965].

b. Handlungsoptionen und -empfehlungen

Um digitaler Exklusion verschiedener, insbesondere auch vulnerabler Gruppen zu begegnen, sind eine Vielzahl von Möglichkeiten erdenklich sowie notwendig.[966] Anhand von fortgesetzter nutzerzentrierter sowie aus Nutzererfahrungen lernender Maßnahmen sollten Nutzerbedürfnisse in einem iterativen Entwicklungsprozess stets besser umgesetzt werden.[967] Um digitaler Exklusion zu begegnen, können unterschiedliche (analoge und digitale) Kommunikationskanäle derart miteinander verbunden werden, sodass ein Zugang weitestgehend aller Bevölkerungsgruppen gewährleistet werden kann, beispielsweise über Telefon- oder Webchat-Assistenz, weiterhin mögliche schriftliche Einreichung von Dokumenten, Online-Zentren – etwa in Bürgerämtern – zur Unterstützung vor Ort als auch Assistenzen durch mobile Helfer

---

960 Siehe hierzu ausführlich *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025.
961 *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 115ff., 242ff., 250f., 265ff., 356ff.; nach einer Erhebung von Eurostat bestehen derzeit bei 46 % der europäischen Bürgerinnen und Bürger keine grundlegenden digitalen Ressourcen oder Kompetenzen, *Stürner*, GPR 2025, 62, 69, mit Hinweis auf *Eurostat*, Niveau der digitalen Kompetenzen von Einzelpersonen (2015–2019), abrufbar unter https://ec.europa.eu/eurostat/databrowser/view/isoc_sk_dskl_i$DV_317/default/table?lang=de; vgl. *Turbék*, Justiz und KI, Friedrich-Naumann-Stiftung für die Freiheit (Hrsg.), 2026, abrufbar unter https://shop.freiheit.org/#!/Publikation/2184 (Abruf am 09.10.2026, 15:40 Uhr), S. 33.
962 *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 115ff.
963 *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 356ff.
964 *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 242ff., 250f.
965 *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 265ff.
966 *Dörr*, Elektronischer Rechtsverkehr und Bürgerzugänge, in: Reuß/Windau (Hrsg.), Göttinger Kolloquien zur Digitalisierung des Zivilverfahrensrechts, Kolloquien im Sommersemester 2021, Göttingen 2022, S. 119.
967 *Dörr*, Elektronischer Rechtsverkehr und Bürgerzugänge, in: Reuß/Windau (Hrsg.), Göttinger Kolloquien zur Digitalisierung des Zivilverfahrensrechts, Kolloquien im Sommersemester 2021, Göttingen 2022, S. 119.

mit mobilen Endgeräten.[968] Auch hierbei können Zentralisierungen vorteilhaft eingesetzt werden. So bietet der Infoservice Niedersächsische Justiz bereits heute eine anschlussfähige Lösung, indem geschultes Personal, das mit einer „mitlernenden“ Software ausgestattet ist, Bürgerinnen und Bürgern zentral Auskunft zu nichtverfahrensbezogenen Fragen erteilt.[969] Die Software wertet dabei die Fragen aus und unterbreitet den Justizmitarbeitern Antwortoptionen; von diesen wählen jene die zutreffende aus und können so den Bürgerinnen und Bürgern qualitative Antworten unterbreiten sowie die Software durch die Auswahl verbessern und zugleich das ansonsten bei den Gerichten in erheblichem Umfang mit der Beantwortung von allgemeinen Fragen befasste Personal entlasten, das sich in der Folge um verfahrensbezogene Aufgaben kümmern kann.[970]

Bei den Gerichten sollten **physische Rechtsantragstellen** zur Minimierung der Zugangsschwellen eingerichtet werden, die Rechtsschutzsuchende bei der Eingabe ihrer Daten unterstützen.[971] Zusätzlich dazu könnten auch „virtuelle Antragstellen“ eingerichtet werden, bei denen mit rechtsschutzsuchenden Personen, die entsprechend notwendige Kompetenzen besitzen, über eine Videokonferenz kommuniziert wird.[972]

Die **Absenkung der Gerichtskosten** sollte für Verfahren erwogen werden, in denen entscheidungsersetzende informationstechnische Systeme eingesetzt werden, jedenfalls sofern dieser Einsatz mit geringeren Kosten einhergeht, als wenn eine menschliche Richterin oder ein menschlicher Richter mit dem Verfahren betraut wird.[973] Hierdurch kann die Attraktivität des gerichtlichen Verfahrens auch gegenüber alternativen Streitbeilegungsmöglichkeiten erhöht und der *rationalen Apathie*[974] entgegengewirkt werden.[975]

Vorbehaltlich der niedrigschwelligen – physischen – Unterstützungsmöglichkeiten beim Zugang zu Gericht, kann eine „bundeseinheitliche Kommunikationsplattform“

---

[968] *Dörr*, Elektronischer Rechtsverkehr und Bürgerzugänge, in: Reuß/Windau (Hrsg.), Göttinger Kolloquien zur Digitalisierung des Zivilverfahrensrechts, Kolloquien im Sommersemester 2021, Göttingen 2022, S. 119.

[969] Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 78f.

[970] Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 78f.

[971] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 116, 251, 357.

[972] Thesenpapier der Arbeitsgruppe „Modernisierung des Zivilprozesses“ im Auftrag der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://www.justiz.bayern.de/media/images/behoerden-und-gerichte/oberlandesgerichte/nuernberg/thesenpapier_der_arbeitsgruppe.pdf, (Abruf am 26.02.2026, 13:01 Uhr).

[973] Vgl. *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 116 m. w. N.

[974] Dies bezeichnet Konstellationen, in denen die voraussichtlichen Kosten eines Verfahrens den Nutzen für den Einzelnen überwiegen. Vgl. hierzu im Bereich des Verbraucherrechts *Weber*, VuR 2013, 323, 325.

[975] Vgl. hierzu *Hoffmann*, RDi 2022, 76, 81f. sowie *Weber*, VuR 2013, 323, 325.

geschaffen werden.[976] Eine transparente Handhabe kann hierbei die Akzeptanz der Bürger unterstützen, wenngleich diese frei entscheiden können sollten, wie sie mit der Justiz kommunizieren.[977] Bei derartigen Online-Eingaben, etwa in Form von Masken, ist mit Blick auf eine größere Zugänglichkeit und Effizienz zu beachten, dass diese nicht zu kompliziert oder umfangreich zum Ausfüllen sind.[978]
Der Zugang zu Verfahren im entscheidungsersetzenden Einsatz, die präsumtiv vornehmlich digital über eine Online-Plattform stattfinden werden, kann zudem durch weitere digitale Barrieren, etwa aufgrund von Behinderungen, erschwert werden. Barrierefreiheitsregelungen, die den staatlichen Bereich bzw. den justiziellen Bereich adressieren, finden sich auf verschiedenen Ebenen. Genannt seien hierfür exemplarisch Art. 9 Abs. 1, 13 Abs. 1 UN-Behindertenrechtskonvention, die Richtlinie (EU) 2016/2102 über den barrierefreien Zugang zu den Websites und mobilen Anwendungen öffentlicher Stellen, §§ 186 und 191a GVG, das bundesrechtliche Behindertengleichstellungsgesetz mit seiner Umsetzungsverordnung, der Barrierefreie-Informationstechnik-Verordnung, sowie die landesrechtlichen Pendants (z. B. Art. 14 Bayerisches Behindertengleichstellungsgesetz mitsamt §§ 9 ff. der Verordnung über die Digitalisierung im Freistaat Bayern).[979] Digitale Barrieren können für alle Menschen auftreten (Bsp.: Infolge eines Armbruchs wird die Bedienung einer Computermaus erschwert/verunmöglicht.) und auch mit Blick auf den demographischen Wandel und die älter werdenden Bürgerinnen und Bürger ist die entsprechend barrierefreie Ausgestaltung digitaler Angebote von hoher Wichtigkeit.[980]
Entsprechend der Vielzahl der Gegebenheiten, die in möglichen digitalen Barrieren resultieren können, besteht eine Vielzahl an Maßnahmenoptionen, selbige abzumildern bzw. zu überwinden. So benötigen Menschen mit Sehbehinderungen z. B. Vergrößerungssoftware bzw. bei Erblindung Braillezeilen und

---

[976] Die *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, Münchener Thesen zum Zivilprozess der Zukunft, abrufbar unter https://www.olg-duesseldorf.nrw.de/infos/Zivilprozess-der-Zukunft/Muenchener-Thesen-zum-Zivilprozess-der-Zukunft.pdf (Abruf am 26.02.2026, 18:27 Uhr), S. 4f., These A.1.

[977] Die *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, Münchener Thesen zum Zivilprozess der Zukunft, abrufbar unter https://www.olg-duesseldorf.nrw.de/infos/Zivilprozess-der-Zukunft/Muenchener-Thesen-zum-Zivilprozess-der-Zukunft.pdf (Abruf am 26.02.2026, 18:27 Uhr), S. 4f., These A.1.

[978] *Meller-Hannich*, AnwBl 2021, 288, 290; hingewiesen sei an dieser Stelle auf die Pilotierung der Einreichungsmöglichkeit von Zahlungsklagen im Online-Verfahren bis zu einem Streitwert von 10.000 €, die am 15.04.2026 an den ersten pilotierenden Amtsgerichten bekanntgegeben wurde, Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 75f.. Die Erfahrungen dieser Pilotierung können vorteilhafterweise im weiteren Digitalisierungsprozess Einzug erhalten.

[979] Siehe beispielhaft auch *E-Justice-Rat*, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr), S. 26.

[980] Vgl. *Bundesministerium für Arbeit und Soziales*, Leitlinien für die Anwendung des Barrierefreiheitsstärkungsgesetzes, abrufbar unter https://www.bundesfachstelle-barrierefreiheit.de/SharedDocs/Downloads/DE/Externe-Veroeffentlichungen/bmas-leitlinien-bfsg (Abruf am 23.09.2026, 20:12 Uhr), S. 2; *Bundesministerium für Digitales und Staatsmodernisierung*, Digitale Barrierefreiheit, Stand: 17.06.2026, abrufbar unter https://www.barrierefreiheit-dienstekonsolidierung.bund.de/Webs/PB/DE/barrierefreie_it/digitale-barrierefreiheit/digitale-barrierefreiheit-node.html (Abruf am 23.09.2026, 20:16 Uhr).

Sprachausgaben.[981] Gehörlose Menschen können hingegen z. B. von visuellen Inhalten und Untertiteln profitieren, während für Menschen mit motorischen Einschränkungen barrierefreie Schnittstellen und die Unterstützung verschiedener Eingabemethoden vorteilhaft sein können.[982] Menschen mit kognitiven Einschränkungen können hingegen beispielsweise durch einfach verständliche Texte in guter Struktur unterstützt werden.[983] Auch KI-Systeme können dazu beitragen, diese Barrieren zu überwinden, z. B. indem sie automatisiert Untertitel erstellen oder Texte in einfache Sprache übersetzen.[984] Da die rechtsprechende Tätigkeit allen Bürgerinnen und Bürgern offensteht, ist auf die umfangreiche Barrierefreiheit des Zugangs zu Gericht – auch und gerade im Fall des entscheidungsersetzenden Einsatzes informationstechnischer Systeme bei der richterlichen Entscheidungsfindung unter Nutzung von Online-Plattformen – maßgebliches Gewicht zu legen und diese sollte daher umfassend umgesetzt werden.

---

[981] *Bundesministerium für Digitales und Staatsmodernisierung*, Digitale Barrierefreiheit, Stand: 17.06.2026, abrufbar unter https://www.barrierefreiheit-dienstekonsolidierung.bund.de/Webs/PB/DE/barrierefreie_it/digitale-barrierefreiheit/digitale-barrierefreiheit-node.html (Abruf am 23.09.2026, 20:16 Uhr) sowie vertiefend und mit Verweis auf weitere Hilfsmittel *Stabsstelle LBIT – Landeskompetenzzentrum für barrierefreie Informationstechnik und Durchsetzungs- und Überwachungsstelle*, Zielgruppe Sehen, abrufbar unter https://lbit.hessen.de/landesbeauftragte-fuer-barrierefreie-it/zielgruppen-barrierefreier-gestaltung/zielgruppe-sehen (Abruf am 23.09.2026, 20:33 Uhr).

[982] *Bundesministerium für Digitales und Staatsmodernisierung*, Digitale Barrierefreiheit, Stand: 17.06.2026, abrufbar unter https://www.barrierefreiheit-dienstekonsolidierung.bund.de/Webs/PB/DE/barrierefreie_it/digitale-barrierefreiheit/digitale-barrierefreiheit-node.html (Abruf am 23.09.2026, 20:16 Uhr) sowie vertiefend und mit Verweis auf weitere Hilfsmittel *Stabsstelle LBIT – Landeskompetenzzentrum für barrierefreie Informationstechnik und Durchsetzungs- und Überwachungsstelle*, Zielgruppe Hören, abrufbar unter https://lbit.hessen.de/landesbeauftragte-fuer-barrierefreie-it/zielgruppen-barrierefreier-gestaltung/zielgruppe-hoeren (Abruf am 23.09.2026, 20:35 Uhr) und *Stabsstelle LBIT – Landeskompetenzzentrum für barrierefreie Informationstechnik und Durchsetzungs- und Überwachungsstelle*, Zielgruppe Verstehen & Interagieren, abrufbar unter https://lbit.hessen.de/landesbeauftragte-fuer-barrierefreie-it/zielgruppen-barrierefreier-gestaltung/zielgruppe-interaktion-verstaendnis (Abruf am 23.09.2026, 20:36 Uhr).

[983] *Bundesministerium für Digitales und Staatsmodernisierung*, Digitale Barrierefreiheit, Stand: 17.06.2026, abrufbar unter https://www.barrierefreiheit-dienstekonsolidierung.bund.de/Webs/PB/DE/barrierefreie_it/digitale-barrierefreiheit/digitale-barrierefreiheit-node.html (Abruf am 23.09.2026, 20:16 Uhr) sowie vertiefend und mit Verweis auf weitere Hilfsmittel *Stabsstelle LBIT – Landeskompetenzzentrum für barrierefreie Informationstechnik und Durchsetzungs- und Überwachungsstelle*, Zielgruppe Verstehen & Interagieren, abrufbar unter https://lbit.hessen.de/landesbeauftragte-fuer-barrierefreie-it/zielgruppen-barrierefreier-gestaltung/zielgruppe-interaktion-verstaendnis (Abruf am 23.09.2026, 20:36 Uhr).

[984] Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 43 m. w. N.; für die KI-gestützte Übersetzung von Texten in Leichte Sprache existiert sogar bereits ein Kodex, *Borchert/Bredowski/Delor/Eppig/Göhner/Jenke/Jessen-Narr/Krimitzas/Lange-Kniep/Nägele/Nietzio/Nolte/Rhensius/Schnepel*, Kodex für den Einsatz von Künstlicher Intelligenz für Leichte-Sprache-Texte, 2025, abrufbar unter https://www.dicv-limburg.de/derverband/wissenswertes/leichte-sprache/kodex-fuer-ki-gestuetzte-uebersetzungen/kodex-fuer-ki-gestuetzte-uebersetzungen (Abruf am 13.10.2026, 20:16 Uhr).

c. Zusammenfassung

Alle Rechtsschutzsuchenden müssen einen angemessenen Zugang zu Gericht erhalten. Findet der Zugang im entscheidungsersetzenden Einsatz vornehmlich digital über eine Online-Plattform statt, sind für Personen ohne die notwendigen digitalen Kompetenzen oder Ressourcen bei den Gerichten niedrigschwellige Unterstützungsmöglichkeiten einzurichten.
Um digitaler Exklusion zu begegnen, sind eine Vielzahl an Möglichkeiten denkbar und notwendig. Anhand von fortgesetzter nutzerzentrierter sowie aus Nutzererfahrungen lernender Maßnahmen sollten Nutzerbedürfnisse in einem iterativen Entwicklungsprozess stets besser umgesetzt werden. Analoge und digitale Kommunikationskanäle sollten derart verbunden werden, dass weitestgehend alle Bevölkerungsgruppen Zugang erhalten. Dies kann beispielsweise über Telefon- oder Webchat-Hilfe geschehen, eine weiterhin bestehende Möglichkeit der schriftlichen Einreichung von Dokumenten oder mobile Helfer. Auch hierbei können Zentralisierungen vorteilhaft eingesetzt werden. Hierbei bietet der Infoservice Niedersächsische Justiz bereits heute eine anschlussfähige Lösung, mit der zugleich das Gerichtspersonal entlastet wird. Die Gerichte sollten physische Rechtsantragstellen einrichten, die bei der Dateneingabe unterstützen; ergänzend könnten virtuelle Antragstellen per Videokonferenz eingerichtet werden. Für Verfahren mit entscheidungsersetzenden Systemen sollte eine Absenkung der Gerichtskosten erwogen werden. Dies kann die Attraktivität des gerichtlichen Verfahrens auch gegenüber alternativen Streitbeilegungsmöglichkeiten erhöhen und rationaler Apathie entgegenwirken. Möglich ist zudem das Hinzutreten einer bundeseinheitlichen Kommunikationsplattform. Der digitale Zugang sollte umfassend barrierefrei ausgestaltet werden, insbesondere für Menschen mit Behinderungen, aber auch mit Blick auf vorübergehende Einschränkungen und den demographischen Wandel. Nötig sind Maßnahmen, die die jeweilige Einschränkung adressieren. Auf die umfangreiche Barrierefreiheit sollte maßgebliches Gewicht gelegt werden.

9. Sicherstellung der Freiwilligkeit der Systemnutzung durch Richterinnen und Richter

a. Rote Linien entscheidungsunterstützender Einsatz

Richtern und Richterinnen steht es frei, sich im Rahmen ihrer Tätigkeit eines informationstechnischen Systems zu bedienen. Eine **Nutzungsverpflichtung** wäre mit der richterlichen Unabhängigkeit gemäß Art. 97 GG[985] und mit Art. 101 Abs. 1 Satz 2 GG[986], Art. 6 Abs. 1 EMRK[987] sowie Art. 47 Abs. 1, Abs. 2 GRCh[988] nicht zu vereinbaren.[989]

b. Handlungsoptionen und -empfehlungen

Während eine Nutzungsverpflichtung (derzeit) mit der richterlichen Unabhängigkeit nicht zu vereinbaren ist, da die Funktionsfähigkeit der Gerichte infolge des Arbeitsanfalls nicht nachhaltig gefährdet ist, kann sich dies zukünftig u. a. infolge des sozio-demographischen Wandels durchaus ändern.[990] Hierbei ist zu beachten, dass eine Beeinträchtigung der richterlichen Unabhängigkeit infolge einer unzureichenden Mittelausstattung eintreten kann,[991] sodass beispielsweise auch das Unterlassen der Bereitstellung von informationstechnischen Systemen einen Verstoß gegen Art. 97 GG darstellen könnte. Um diesem Szenario vorzubeugen, sollte **bereits frühzeitig die Implementierung informationstechnischer Systeme** im Rahmen der richterlichen Entscheidungsfindung angestrebt werden und die Entwicklung dieses Feldes eng sowie aktuell begleitet werden. Hierbei sollte die Nutzung informationstechnischer Systeme bereits als Langzeitprojekt verstanden werden und eine hinreichende Mittelausstattung für die Gerichte eingeplant werden; die Nahelegung der Nutzung durch Justizverwaltungen wird hingegen bereits für unzulässig gehalten.[992]

---

[985] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 192f., 199, 209.
[986] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 221ff.
[987] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 245f.
[988] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 265ff.
[989] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 192f., 199ff., 209, 221ff., 245f., 265ff.
[990] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 192f., 199, 209.
[991] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 191f. m. w. N.
[992] *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 283; *Arbeitsgruppe „Einsatz von KI und algorithmischen Systemen in der Justiz“*, Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier zur 74. Jahrestagung der Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs vom 23. bis 25. Mai 2022 in Rostock, abrufbar unter https://www.justiz.bayern.de/media/images/behoerden-und-gerichte/oberlandesgerichte/nuernberg/einsatz_von_ki_und_algorithmischen_systemen_in_der_justiz.pdf (Abruf am 15.04.2026, 18:58 Uhr), S. 12; Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 13, 127f.

Es ist ferner sicherzustellen, dass infolge der effizienteren Erledigung gerichtlicher Verfahren nicht auch die Zielvorgaben im Rahmen der **Personalbedarfsbemessung (PEBB§Y)** erhöht werden, da dies zu einem faktischen Nutzungszwang führen würde, der entsprechend die sachliche sowie die persönliche richterliche Unabhängigkeit beeinträchtigte.[993] Dies ist auch bei einer Neugestaltung des PEBB§Y-Systems bzw. eines entsprechenden Systems zu beachten.[994]
Auch sind faktische, kritisch anzusehende Nutzungsverpflichtungen im Zusammenhang mit **dienstlichen Beurteilungen** auszuschließen.[995] Eine Berücksichtigung der Nutzung digitaler Arbeitsmittel (umfassend informationstechnische Systeme bei der richterlichen Entscheidungsfindung) kann jenseits hiervon jedoch dem Grunde nach Einzug in die dienstliche Beurteilung halten.[996] Im Zuge der Nutzung von informationstechnischen Systemen bei der richterlichen Entscheidungsfindung ist hierbei sicherzustellen, dass trotz der Nutzung dieser weiterhin hinreichende Feststellungen zu der Fachkompetenz sowie zu den Fachkenntnissen getroffen werden können müssen.[997] Die Beurteilungsrichtlinien sowie Beurteilungsvordrucke der Dienststellen sind entsprechend anzupassen und

---

[993] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 194f.; *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 286; *Arbeitsgruppe „Einsatz von KI und algorithmischen Systemen in der Justiz“*, Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier zur 74. Jahrestagung der Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs vom 23. bis 25. Mai 2022 in Rostock, abrufbar unter https://www.justiz.bayern.de/media/images/behoerden-und-gerichte/oberlandesgerichte/nuernberg/einsatz_von_ki_und_algorithmischen_systemen_in_der_justiz.pdf (Abruf am 15.04.2026, 18:58 Uhr), S. 12f.; siehe auch *Dötterl*, RDi 2026, 414, 419 sowie *Turbék*, Justiz und KI, Friedrich-Naumann-Stiftung für die Freiheit (Hrsg.), 2026, abrufbar unter https://shop.freiheit.org/#!/Publikation/2184 (Abruf am 09.10.2026, 15:40 Uhr), S. 16f., 33; siehe auch Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 13.
[994] *Gold/Reuß et al.*, Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. 195 m. w. N.
[995] *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 287; *Arbeitsgruppe „Einsatz von KI und algorithmischen Systemen in der Justiz“*, Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier zur 74. Jahrestagung der Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs vom 23. bis 25. Mai 2022 in Rostock, abrufbar unter https://www.justiz.bayern.de/media/images/behoerden-und-gerichte/oberlandesgerichte/nuernberg/einsatz_von_ki_und_algorithmischen_systemen_in_der_justiz.pdf (Abruf am 15.04.2026, 18:58 Uhr), S. 12f.
[996] *Biallaß* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel 8.1 (Stand: 01.09.2025), Rn. 287; *Arbeitsgruppe „Einsatz von KI und algorithmischen Systemen in der Justiz“*, Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier zur 74. Jahrestagung der Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs vom 23. bis 25. Mai 2022 in Rostock, abrufbar unter https://www.justiz.bayern.de/media/images/behoerden-und-gerichte/oberlandesgerichte/nuernberg/einsatz_von_ki_und_algorithmischen_systemen_in_der_justiz.pdf (Abruf am 15.04.2026, 18:58 Uhr), S. 12f.
[997] Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026 der *Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs*, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr), S. 37.

sollten entsprechende ausdrückliche Klarstellungen zu diesen Thematiken erhalten. Auf die Einhaltung der rechtlichen Grenzen ist seitens der Dienststelle zu achten.

c. Zusammenfassung

Richtern und Richterinnen steht es frei, sich im Rahmen ihrer Tätigkeit eines informationstechnischen Systems zu bedienen. Eine Nutzungsverpflichtung ist (derzeit) mit der richterlichen Unabhängigkeit unvereinbar.
Diese derzeitige Unvereinbarkeit der Nutzungsverpflichtung mit der richterlichen Unabhängigkeit kann sich jedoch u. a. durch den sozio-demographischen Wandel ändern, wenn die Gerichte in ihrer Funktionsfähigkeit nachhaltig infolge des Arbeitsanfalls gefährdet sind. Zu beachten ist, dass auch eine unzureichende Mittelausstattung die richterliche Unabhängigkeit beeinträchtigen kann, z. B. infolge des Unterlassens der Bereitstellung informationstechnischer Systeme. Daher ist eine bereits frühzeitige Implementierung informationstechnischer Systeme sowie die enge und aktuelle Begleitung der Entwicklung dieses Feldes zu empfehlen. Die Nutzung sollte als Langzeitprojekt verstanden und eine hinreichende Mittelausstattung der Gerichte eingeplant werden. Die Nahelegung der Nutzung durch die Justizverwaltung wird hingegen bereits für unzulässig erachtet. Ein faktischer Nutzungszwang, z. B. durch die Erhöhung der Zielvorgaben im Rahmen der Personalbedarfsmessung, ist auch bei der Neugestaltung des entsprechenden Systems zu verhindern. Gleiches gilt für einen faktischen Nutzungszwang im Zusammenhang mit dienstlichen Beurteilungen. Beurteilungsrichtlinien und -vordrucke sind anzupassen und insbesondere mit Klarstellungen hierzu zu versehen.

10. Gesamtzusammenfassung

a. Untersuchungsansatz und Leitprinzip der Handlungsempfehlungen

Das vorliegende Arbeitspaket (Milestone 2 & 3) knüpft an die im ersten Arbeitspaket (Milestone 1) des Forschungsprojekts „Künstliche Intelligenz und richterliche Entscheidungsfindung“ herausgearbeiteten rechtlichen Rahmenbedingungen aus dem Verfassungs-, Menschen-, Europa- und Zivilprozessrecht für den Einsatz informationstechnischer Systeme bei der richterlichen Entscheidungsfindung an. Es aggregiert die in den einzelnen Untersuchungsgegenständen identifizierten rechtlichen Grenzen zu roten Linien und ordnet sie in acht Themenfelder: Menschliche Aufsicht und Kontrolle des Systems, Schutz informationstechnischer Systeme vor Störungen, Transparenz staatlichen Handelns und Information über den Einsatz von Künstlicher Intelligenz, Datenschutz, Notwendigkeit einer Rechtsgrundlage, Verfügbarkeit von Rechtsbehelfen, Barrierefreiheit des Rechtsschutzes sowie Sicherstellung der Freiwilligkeit der Systemnutzung durch Richterinnen und Richter. Hierbei unterscheidet das vorliegende Arbeitspaket (Milestone 2 & 3) zwischen übergreifenden roten Linien, die für jeden Einsatz gelten, und roten Linien, die nur den entscheidungsersetzenden oder nur den entscheidungsunterstützenden Einsatz betreffen. Aufbauend auf dieser Verdichtung und Systematisierung der rechtlichen Grenzen für den Einsatz informationstechnischer Systeme bei der richterlichen Entscheidungsfindung werden den jeweiligen roten Linien Handlungsoptionen zugeordnet, die einen diese Grenzen wahrenden Einsatz ermöglichen sowie Handlungsempfehlungen im Sinne eines „Gold-Standards“ herausgearbeitet für die bestmögliche Verwirklichung der konkreten rechtlichen Anforderung.
Das leitende Prinzip vieler dieser Empfehlungen ist die im ersten Arbeitspaket (Milestone 1) entwickelte dreigliedrige Funktionsmatrix, bestehend aus dem funktionalen Charakter der Tätigkeit, dem Grad der Automatisierung sowie der Komplexität des Anwendungsfeldes. Je stärker ein System sachlich-inhaltliche Aufgaben wahrnimmt, je höher der Automatisierungsgrad ist und je komplexer das Anwendungsfeld gestaltet ist, desto höher sind die Anforderungen zur Einhaltung der ethischen Einsatzgrenzen und der rechtlichen Rahmenbedingungen. Das vorliegende Arbeitspaket (Milestone 2 & 3) zieht hieraus insbesondere die Konsequenz, dass das zu implementierende Governance-Modell, die Ansiedelung der Entwicklung und des Betriebs des informationstechnischen Systems sowie das Modell für die Ausfüllung der richterlichen Kontrollpflicht nicht pauschal, sondern anhand der Funktionsmatrix abgestuft auszugestalten sind.

b. Übergreifende rote Linien, Handlungsoptionen und -empfehlungen

Die menschliche Aufsicht und Kontrolle ist eine der zentralen roten Linien für jeden Einsatz informationstechnischer Systeme bei der richterlichen Entscheidungsfindung. Richterinnen und Richter sind umfassend in die Beschaffung, den Betrieb und die Überwachung des Systems einzubinden. Beim entscheidungsersetzenden Einsatz und bei entscheidungsunterstützenden Funktionen, die den Kernbereich der Rechtsprechung berühren, sind sie entsprechend auch in die Entwicklung einzubinden. Die Beauftragung Privater oder der Verwaltung mit der Entwicklung ist möglich, sofern der hinreichende richterliche Einfluss gewahrt ist. Eine inhaltliche Kontrolle richterlicher Dokumente durch Dritte ist auszuschließen. Unbedingt anzuraten ist ein richterliches Governance-Modell, das die Aufsicht und Kontrolle über die identifizierten Bereiche institutionell absichert. In Betracht kommen interdisziplinär

besetzte, richterlich(dominiert)e „Automatisierungskammern“, eine auch bundesländerübergreifende Automatisierungsinstitution, deren Kombination oder eine Verbundarbeit mit Dritten (Exekutive oder [private] Dritte). Ein pauschal für alle Einsatzszenarien angemessenes Modell besteht nicht. Vielmehr ist das passende Governance-Modell anhand des konkret geplanten Einsatzes adäquat mittels der Funktionsmatrix zu entwickeln und zu etablieren. Für den derzeit nur in hinreichend einfach gelagerten Fällen möglichen entscheidungsersetzenden Einsatz mittels regelbasierter Systeme ist z. B. eine richterliche Beteiligung im Rahmen eines interdisziplinären Entwicklungsteams denkbar, wobei die v. a. inhaltliche Überwachung an zuvor im Geschäftsverteilungsplan bestimmte Richterinnen und Richter zugewiesen werden kann. Bei rein administrativen Unterstützungsfunktionen dürfte die Überlassung der inhaltlichen Überwachung an den jeweiligen erkennenden Richter genügen, während sich für Funktionen am anderen Ende der Bandbreite ein dauerhaftes Automatisierungsgremium anbieten kann.
Die Ansiedelung der Entwicklung und des Betriebs des informationstechnischen Systems bei der Exekutive oder bei privaten Dritten ist nicht ausgeschlossen, unterliegt jedoch strengen Anforderungen. Zu empfehlen sind in diesem Falle klare, schriftlich fixierte Regelungen der Zugriffsrechte auf das System samt weitreichender Auflagen, deren Einhaltung unter Beteiligung der Richterschaft überwacht wird. Der Judikative sind effektive Kontroll- und Mitsprachemöglichkeiten einzuräumen, wobei die Übertragung der Wahrnehmung dieser Aufgaben an einen richterlichen IT-Beauftragten vorzugswürdig erscheint. Ist eine hinreichende Aufsicht und Kontrolle nicht sicherzustellen, scheidet die externe Ansiedelung aus. Zum Ausschluss einer inhaltlichen Kontrolle richterlicher Dokumente empfiehlt sich eine weitestgehende Trennung technischer und organisatorischer Art der externen Zugriffsrechte sowie richterlichen Arbeitssphäre; die Verwertung der Inhalts- sowie Protokolldaten für dienstaufsichtsrechtliche Zwecke ist grundsätzlich auszuschließen, etwaige Ausnahmen sind klar zu fassen, und schon der „böse Schein“ eines unbefugten Zugriffs auf gerichtliche bzw. richterliche Dokumente muss durch Schutzvorkehrungen ausgeschlossen sein. Die Handlungsempfehlung hinsichtlich der Ansiedelung der Entwicklung und des Betriebs des Systems besteht darin, dass sie umso eher bei der Judikative selbst erfolgen sollte, je stärker das System inhaltsbezogene Aufgaben wahrnimmt. Die vollständige Ansiedelung in judikativer Hand bildet den „Königsweg“. Dies steht auch mit dem Prinzip der technologischen Souveränität im Einklang.
Wirksame Überwachung und Kontrolle setzt die entsprechende Befähigung bzw. Kompetenz hierzu voraus. Neben den rechtlichen sollten auch die notwendigen technischen Kenntnisse dringend vermittelt werden. Dies sollte bedarfsgerecht, möglichst frühzeitig (bereits in Studium und Ausbildung) und in einer mehrstufigen Ausgestaltung erfolgen, ergänzt durch Leitfäden und ggf. verpflichtende Fortbildungsveranstaltungen. Der notwendigen Aufsichts- und Kontrolltätigkeit sollte ein angemessener Stellenwert eingeräumt werden. Sinnvoll erscheint zudem die Implementierung einer Pflicht, systemische Auffälligkeiten, unplausible Ergebnisse oder erkennbare Verzerrungen an eine zentrale Stelle zu melden, eingebettet in ein breit angelegtes Informationsmanagementsystem. Ergänzend kommt eine mehrstufige Qualitätskontrolle in Betracht. Dienstrechtliche Maßnahmen für offenkundige Verletzungen der Aufsichts- und Kontrollpflicht sollten vorgesehen, ihre Voraussetzungen und die jeweiligen Sanktionen jedoch so bestimmt gefasst und kommuniziert werden, dass kein, die Systemnutzung faktisch unterbindender, abschreckender Effekt eintritt.
Die zweite wesentliche Rahmenbedingung im Zuge der menschlichen Aufsicht und Kontrolle des Systems betrifft den Ausschluss von Bias. Ein Einsatz

informationstechnischer Systeme scheidet aus, wenn ein ergebnisrelevanter Bias nicht ausgeschlossen werden kann bzw. die menschliche Kontrolle nicht gewährleistet ist. Eine einheitliche Lösung gegen systemseitige Diskriminierungen und Bias existiert nicht; die Problematik sollte daher auf allen Ebenen der Technik-Architektur adressiert werden, u. a. durch Qualitätskontrollen bereits bei der Systementwicklung, eine diskriminierungssensible Auswahl der Trainingsdaten, ein dokumentiertes Datenmanagement, Kontrollmechanismen auch gegen mittelbare Diskriminierungen und eine fortlaufende Daten- und Ergebnisevaluation. Insbesondere Datenschutz und Antidiskriminierung sollten dabei in einen Ausgleich gebracht werden. Als Trainingsdaten sind möglichst anonymisierte bzw. nicht personenbezogene Daten zu verwenden. Sofern KI-Modelle Falldaten auch nach Eintritt eines Löschungsgrundes benötigen, sollte eine echte flächendeckende Anonymisierung anstatt einer Löschung erfolgen, wofür sich eine Fortentwicklung des Anonymisierungstools ALeKS anbietet. Zudem sollte eine gesetzliche Pflicht zur Veröffentlichung begründeter gerichtlicher Entscheidungen eingeführt werden.
Eine weitere übergreifende rote Linie ist der Schutz informationstechnischer Systeme vor internen und externen Störungen sowie vor unzulässigem Einfluss Dritter. Die Sicherheit des informationstechnischen Systems ist über seinen gesamten Lebenszyklus zu betrachten, wofür Risikobewertungen für die jeweiligen Systeme, ein angepasstes Informationssicherheitsmanagementsystem sowie ein präventives, umfassendes Sicherheits- und Cyber-Resilienzkonzept relevant sind. Wesentliches Augenmerk sollten präventive Absicherungen mit umfangreichen Testphasen und Evaluierungen vor dem Ausrollen der Systeme erfahren. Weitestmögliche Vorkehrungen gegen Prompt-Injection-Angriffe sollten ergriffen werden. Unabhängig von der Einordnungsfrage von KI-Systemen als Hochrisiko-KI-System sollte stets ein Risikomanagementsystem nach Art. 9 KI-VO als kontinuierlicher, iterativer Prozess eingerichtet werden. Jedenfalls sollte zumindest eine Bund/Länder-übergreifende Verständigung auf Expertengrundlage zur Auslegung der KI-VO erarbeitet und etabliert werden. Risiken sollten möglichst unmittelbar technisch aufgelöst werden nach dem Grundsatz „Konstruktion vor Instruktion“. Systemtests sind vor der Inbetriebnahme durchzuführen und regelmäßig zu wiederholen, u. a. unter Nutzung von Reallaboren. Für den judikativen Einsatz ist höchste Genauigkeit eine elementare Voraussetzung. Für die Robustheit erscheinen insbesondere technische Redundanzen als wesentlich. Mit Blick auf die Cybersicherheit sind insbesondere KI-spezifische Angriffstaktiken zu beachten.
Die KI-Lieferkette ist bei Bezug informationstechnischer Systeme von Dritten als eigener Risikofaktor zu behandeln. Als Handlungsoption wird vorgeschlagen, die unmittelbaren Bezugspartner zu verpflichten, die von weiteren Dritten bezogenen Bestandteile auf die für die Justiz notwendigen Sicherheitsstandards zu überprüfen und hierüber regelmäßig Bericht zu erstatten. Die Manipulationsfreiheit sollte durch externe vertrauenswürdige Institutionen bzw. justizeigene Expertengremien überprüft werden. Zu Zertifizierungen sollte eine fortlaufende Auditierung über den gesamten Lebenszyklus hinzutreten. Die Zertifizierung sollte auch die Diskriminierungsfreiheit und die Abbildung sämtlicher Auslegungsvarianten umfassen. Ist die Rechtsperson, die das System entwickelt oder betreibt, selbst Partei eines zu entscheidenden Rechtsstreits oder hat sie am Verfahrensausgang (offenkundig) eigene Interessen, sollte angedacht werden, das System nicht zu nutzen. Dies ist insbesondere bei der Ansiedelung des Systems in der Hand der Exekutive relevant im Hinblick auf Gerichtsbarkeiten, bei denen der Staat regelmäßig Prozessbeteiligter ist. Hier bestätigt sich hervorgehoben die Ansiedelung bei der Judikative als „Königsweg“.

Mit Blick auf den Datenschutz bietet § 3 BDSG grundsätzlich für beide Einsatzformen eine Rechtsgrundlage, für den entscheidungsersetzenden Einsatz bedarf es hierbei jedoch noch entsprechender Aufgabenzuordnungs- bzw. Zuständigkeitsregelungen und einer rechtlichen Grundlage (i. S. v. Art. 22 Abs. 2 lit. b DSGVO) zur Überwindung des Verbots automatisierter Entscheidungsfindungen nach Art. 22 Abs. 1 DSGVO. Als rote Linien sind weiterhin die Gewährleistung von Datensicherheit, die Löschbarkeit personenbezogener Daten, ohne die ein Einsatz grundsätzlich unzulässig ist, Informationspflichten sowie Berichtigungsrechte und die regelmäßige Unzulässigkeit, personenbezogene Daten in Systeme einzugeben, die nicht der Kontrolle der Justiz unterstehen bzw. die die datenschutzrechtlichen Anforderungen nicht wahren, zu verzeichnen. Unter dem Eindruck, dass der breite Einsatz informationstechnischer Systeme im Justizsystem bereits gelebte Praxis ist, wird empfohlen, dieses Verbot und die Folgen eines Verstoßes gegen dieses deutlich zu kommunizieren und gleichzeitig rechtmäßige Alternativen anzubieten. Hinsichtlich der nach Art. 24 DSGVO zu ergreifenden Maßnahmen existiert kein verbindlicher Mindeststandard. Eine Orientierung an Katalogen mit Referenzmaßnahmen ist hingegen möglich und sollte erfolgen. Die Anforderungen des Art. 25 DSGVO sind bereits in der Konzeptphase im Eigenentwicklungsfall bzw. bei der Auswahl externer Dritter für die Entwicklung zu beachten. Weil die Datenschutzaufsichtsbehörden nicht für die Aufsicht über Verarbeitungen zuständig sind, die Gerichte im Rahmen ihrer justitiellen Tätigkeit vornehmen und eine justitielle Selbstkontrolle bislang nicht umgesetzt ist, sollte die Schaffung einer die gesamte richterliche Sphäre abdeckenden Aufsichtsstruktur in Betracht gezogen werden. Die für die richterliche Entscheidungsfindung relevanten Verarbeitungen sollten in eine Muss-Liste nach Art. 35 Abs. 4 DSGVO für die Datenschutz-Folgeabschätzung aufgenommen werden. Hinsichtlich der Sicherstellung von Betroffenenrechten ist nicht deren Beschränkung auf ein vertretbares Minimum, sondern die weitgehende Zulassung oder Stärkung der datenschutzrechtlichen Ansprüche zu empfehlen. Mit Blick auf die Datenverarbeitungs-Informationspflichten ist insbesondere eine einheitliche, mit der Bereitstellung anderer notwendiger Informationen abgestimmte Informationsstrategie zu empfehlen.

Die Transparenzanforderungen staatlichen Handelns verlangen für jeden Einsatz die Verfahrensparteien und die Öffentlichkeit unaufgefordert über das „Ob“ und das „Wie“ des Einsatzes eines informationstechnischen Systems zu informieren und die abstrakte Funktionalität transparent zu machen. Die Parteien müssen die involvierte Logik des Systems nachvollziehen und die wesentlichen Gründe erkennen können, auf denen die Entscheidung des Systems beruht. Eine Offenlegung des Quellcodes ist grundsätzlich nicht erforderlich. Die Black-Box-Problematik steht der Erfüllung einer Transparenz- und Informationspflicht nicht entgegen, da eine nachträgliche Entscheidungsbegründung dem Transparenzgebot hinreichend Rechnung trägt. Die Mitteilung des Systemeinsatzes sollte vor bzw. spätestens zu Verfahrensbeginn erfolgen und über die Art des Systems sowie seine abstrakte Funktionslogik informieren. Einer „Desinformation durch Information“ lässt sich durch ein abgestuftes, verständliches und barrierefreies Informationsmodell begegnen. Die Transparenzpflichten nach Art. 50 KI-VO sollten unmittelbar in die prozessuale Verfahrens- und Arbeitskommunikation eingebettet werden. Zur Konkretisierung der Pflichten nach Art. 50 ff. KI-VO bietet es sich insbesondere an, Leitlinien, Praxisleitfäden und Verhaltenskodizes zu verfolgen und regelmäßig auf vorteilhafte Maßnahmen zu überprüfen.

Für die Überprüfbarkeit des Systems ist ein mehrstufiges Informationsmodell eine wesentliche Handlungsoption. Eine tiefgehende Kontrolle setzt letztlich den Zugang

von Spezialisten zum System und seinen Bestandteilen voraus. Eine proaktive Veröffentlichung des Quellcodes könnte zwar vertrauensbildend wirken, kollidiert jedoch mit Sicherheitsinteressen und sollte daher nicht umgesetzt werden. Zu befürworten ist eine restriktive Handhabung der Quellcode-Offenlegung. Diese kann ausnahmsweise notwendig sein, wenn das Offenlegungsinteresse gegenläufige Interessen überwiegt. Dies kann im Rechtsmittelverfahren relevant werden, bei berechtigten Zweifeln an der akkuraten Funktion eines entscheidungsersetzenden Systems. In diesem Fall kann der Quellcode einem Gerichtsgutachter zu Beweiszwecken zugänglich gemacht werden. Diese Konstellationen sollten vorab bedacht und von mit mehreren Richterinnen und Richtern besetzten Spruchkörpern oder einer unabhängigen sachkundigen Institution entschieden werden. Die technische Protokollierung nach Art. 12 KI-VO sollte so genau sein, dass möglichst jede Entscheidung des Systems erklärbar ist. Die Aufbewahrungspflicht trifft in der Regel die das System im eigenen IT-System betreibende Justiz. Die Aufbewahrungsdauer sollte mit den Fristen aus der JAktAV synchronisiert werden.

c. Entscheidungsunterstützender Einsatz

Auch beim entscheidungsunterstützenden Einsatz, der auch mit Systemen maschinellen Lernens in Betracht kommt, können die Systeme die Einhaltung der rechtlichen Anforderungen derzeit nur in hinreichend einfach gelagerten Fällen sicherstellen. Eine maßgebliche rote Linie ist das zwingende Erfordernis richterlicher Kontrolle und ggf. Korrektur. Die Richterin oder der Richter trägt die Letztverantwortung für die Gesetzeskonformität der Entscheidung sowie insbesondere dafür, dass alle verfahrensrelevanten Informationen zutreffend erfasst und verarbeitet wurden und sich ein systemseitiger Bias und systemseitige Halluzinationen nicht ergebnisrelevant niederschlagen. Der Richterin bzw. dem Richter müssen substanzielle Kontroll- und Korrekturmöglichkeiten eingeräumt werden. Psychologische Effekte, wie der Automation Bias, sind durch geeignete systemseitige oder sonstige Vorkehrungen möglichst auszuschließen. Richterinnen und Richter sind über die Funktionsweise und die Limitierungen des Systems zu informieren. Für die Wahrnehmung der bereits aus der Gesetzesbindung nach Art. 20 Abs. 3 GG folgenden Kontrollpflicht ist ein Kontrollmodell zu empfehlen, das abhängig von der Komplexität des Anwendungsfeldes, dem Grad der Automatisierung und der Funktionalität des Systems gestaltet ist. Entsprechend dem Spektrum der Matrix können auch die entsprechend notwendig anzusetzenden Maßnahmen verschieden intensiv sein. So können beispielsweise bei rein administrativen Funktionen niedrigschwelligere Prüf- und Korrekturmöglichkeiten genügen, während für weitreichendere Funktionen zwingende bzw. ausdrückliche Freigaben von (einzelnen) Ausgabeteilen vorgesehen werden können. Denkbar sowie empfehlenswert kann hierbei sein, dass Richterinnen und Richter vor Freigabe der Signaturmöglichkeit zunächst einen bestimmten, stichprobenmäßigen Anteil (z. B. 25–33 %) der verarbeiteten Systemergebnisse überprüfen müssen, was zwar nicht die gebotene vollständige Prüfung ersetzt, jedoch einem Technikübervertrauen einen Gegenimpuls entgegensetzt. Bewusst eingebaute Fehler wären zwar zulässig, sind aus Akzeptanzgründen jedoch nicht zu empfehlen.
Allgemein sollte eine möglichst weitgehende Nachvollziehbarkeit für die Richterinnen und Richter ermöglicht werden hinsichtlich der Kriterien und Aspekte für die Ergebnisgenerierung. Empfohlen wird insbesondere eine herkunftssensitive Benutzeroberfläche. Die Informationen sollten möglichst im Originalwortlaut wiedergegeben werden und die Originalquellen direkt über die Benutzeroberfläche per

Klick im Volltext mit hervorgehobener Fundstelle abrufbar gemacht werden. Korrekturen sollten ebenso niedrigschwellig ermöglicht werden. Systembrüche sollten vermieden und das Assistenzsystem idealerweise in die E-Akte integriert werden. Sinngemäße Zusammenfassungen sind nicht gänzlich abzulehnen, bringen aber weitergehende Kontrollerfordernisse mit sich. Von Scores, welche die Sicherheit einer Ausgabe angeben, ist abzuraten. Diese sind ihrerseits nicht verlässliche Systemausgaben und können durch eine fälschlich hohe Sicherheitsangabe zum Verzicht auf eine tatsächlich gebotene Prüfung verleiten. Halluzinationen bilden – anders als beim entscheidungsersetzenden Einsatz – wegen des richterlichen Kontrollerfordernisses und der richterlichen Letztverantwortung keine den Einsatz ausschließende rote Linie. Sie sollten jedoch weitestmöglich vermieden werden, um Rationalisierungseffekte nicht zu beeinträchtigen. Befürwortet wird zudem eine strukturierte Kommunikation samt einer möglichen Begrenzung des Vortragsumfangs. Hinsichtlich der Sicherstellung von Kontrolle und Aufsicht im Hinblick auf psychologische Effekte empfehlen sich Aufklärung und Schulung, gegebenenfalls in Form periodischer, verpflichtender Fortbildungsvorgaben sowie eine sensibilisierte Benutzeroberflächengestaltung. Zudem sind nach der Funktionsmatrix abgestufte Begründungslasten für die Übernahme von KI-Inhalten in diesem Zusammenhang eine mögliche Vorgehensweise.
Richterinnen und Richter sind adressatengerecht über die Funktionsweise und Limitierungen des Systems zu informieren. Sie benötigen vor allem eine funktionale, verständliche Erläuterung. Schulungen und weitere KI-Kompetenzformate sowie die unmittelbare Abrufbarkeit derartiger Informationen im System erscheinen vorteilhaft. Entwicklungs- und Überwachungsgruppen benötigen einen deutlich höheren Grad an Erklärungstiefe, für die Methoden der „Explainable AI“ Ansätze, bei komplexen Modellen derzeit aber wohl noch keine vollständigen Lösungen bieten.
Mit Blick auf die innere Transparenz ist sicherzustellen, dass KI-unterstützt erstellte Entscheidungsgründe die tatsächlich tragenden Gründe wiedergeben. Hierbei kommt eine systemseitige Strukturvorgabe der Systemausgabe in Betracht, nach der das System seine tragenden Gründe adressatengerecht aufbereitet und dabei auch entgegenstehende Gründe samt deren Entkräftung darstellt. Für Hochrisiko-KI-Systeme verlangt Art. 13 KI-VO keine vollständige Erklärbarkeit oder Veröffentlichung des Quellcodes, jedoch hinreichende, funktionsgerechte Transparenz sowie die Erstellung und Bereitstellung einer Betriebsanleitung. Diese sollte möglichst digital in die Systemoberfläche integriert werden.
Gegenüber den Parteien besteht beim entscheidungsunterstützenden Einsatz mangels direkter Interaktion mit dem KI-System keine Offenlegungspflicht hinsichtlich der KI-Interaktion. Gegenüber den nutzenden Richterinnen und Richtern kann die Interaktion mit einem KI-System offengelegt werden, was jedoch regelmäßig entbehrlich ist. Künstlich hergestellte Inhalte sind nach Art. 50 Abs. 2 KI-VO grundsätzlich menschen- und maschinenlesbar als KI-generierte Inhalte zu kennzeichnen. Die Ausnahmen von der Kennzeichnungspflicht des Art. 50 Abs. 2 S. 3 KI-VO können etwa bei administrativen Funktionalitäten in Betracht kommen, sind jedoch nicht für weitergehende sachlich-inhaltliche Funktionalitäten einschlägig. Die Kennzeichnungspflicht des Anbieters nach Art. 50 Abs. 2 KI-VO geht nicht weiter als die Offenlegungspflicht des Betreibers nach Art. 50 Abs. 4 KI-VO, im Fall der Weiterverarbeitung „synthetisch“ generierten Texts durch Nutzerinnen oder Nutzer, beispielsweise bei der Generierung eines Textteils für ein Urteil durch Richterinnen oder Richter, wenn dieser im finalen Urteil konkret angepasst und weiterverarbeitet worden ist. Im Rahmen des entscheidungsunterstützenden Einsatzes besteht für die Gerichtsentscheidungen zwar die Offenlegungspflicht nach

Art. 50 Abs. 4 UAbs. 2 KI-VO dem Grunde nach, diese entfällt jedoch meist ausnahmeregelungsbedingt durch die richterliche Endkontrolle und Verantwortungsübernahme. Mit Blick auf die von der KI-VO unberührten verfassungsrechtlichen Transparenzerfordernisse ist eine Information über das „Ob“ und das „Wie“ des KI-Einsatzes allerdings auch hier zu empfehlen.
Jedenfalls für den Einsatz inhaltsbezogener unterstützender Funktionen bedarf es aufgrund ihrer stärkeren Grundrechtsrelevanz einer gesonderten gesetzlichen Grundlage für das „Ob“ und „Wie“ des Einsatzes. Diese sollte zeitnah geschaffen werden. Hierbei wird eine bundeseinheitliche Regelung im GVG befürwortet, die ausdrücklich festhält, dass inhaltsbezogene Funktionen zugelassen sind, sofern die notwendigen Anforderungen erfüllt sind. Die Einsatzbereiche und Funktionalitäten des Systems sollten in der Rechtsgrundlage abstrakt-generell und technikoffen umschrieben sein. Klarstellenderweise sollten auch administrative Funktionen erfasst werden. Die Zulässigkeit inhaltsbezogener Funktionen könnte an abstrakte Anforderungen geknüpft werden, deren elementare Grundlagen im Parlamentsgesetz niederzulegen sind und die im Übrigen mittels einer aufgrund einer Verordnungsermächtigung zugunsten des Bundesjustizministeriums geschaffenen Rechtsverordnung einheitlich definiert werden könnten. Die Richterschaft sollte bei der Beschaffung der Systeme eng eingebunden sein. Einer weitergehenden verfahrensrechtlichen Anpassung bedürfte es darüber hinaus nicht, da die Systeme von Richterinnen und Richtern in dem bereits jetzt bestehenden Zuständigkeits- und Organisationssystem genutzt würden. Einer gesonderten datenschutzrechtlichen Rechtsgrundlage bedarf es für den entscheidungsunterstützenden Einsatz angesichts des § 3 BDSG ebenfalls nicht. Da die Letztverantwortung bei der menschlichen Richterin bzw. dem menschlichen Richter verbleibt, ergeben sich auch im Hinblick auf das Rechtsmittelrecht keine Handlungsnotwendigkeiten.
Richtern und Richterinnen steht es frei, sich im Rahmen ihrer Tätigkeit eines informationstechnischen Systems zu bedienen. Eine Nutzungsverpflichtung wäre mit der richterlichen Unabhängigkeit derzeit unvereinbar, was sich bei einer nachhaltigen Gefährdung der Funktionsfähigkeit der Gerichte, etwa infolge des sozio-demographischen Wandels, ändern kann. Auch ein faktischer Nutzungszwang, z. B. durch die Erhöhung der Zielvorgaben im Rahmen der Personalbedarfsmessung oder im Zusammenhang mit dienstlichen Beurteilungen, ist auszuschließen. Da umgekehrt auch eine unzureichende Mittelausstattung, z. B. infolge des Unterlassens der Bereitstellung informationstechnischer Systeme, die richterliche Unabhängigkeit beeinträchtigen kann, sollte die Implementierung frühzeitig angestrebt, als Langzeitprojekt verstanden und mit hinreichenden Mitteln unterfüttert werden.

d. Entscheidungsersetzender Einsatz

Ein entscheidungsersetzender Einsatz ist derzeit nur in hinreichend einfach gelagerten Fällen zulässig. Hierunter zu verstehen sind Fälle mit überschaubaren Tatsachen- und Rechtsfragen sowie einem geringen Auslegungsspielraum, weswegen der Weg, auf dem das informationstechnische System zu seiner Entscheidung gelangt, einer mathematisch-logischen Rechenoperation gleichkommt. Insbesondere bei komplexen Abwägungsentscheidungen scheidet diese Einsatzform aus. Daher muss eine systemseitige Aussteuerung an einen menschlichen Richter bzw. eine menschliche Richterin jederzeit, auch auf Antrag, möglich sein, sobald sich zeigt, dass ein solcher hinreichend einfach gelagerter Fall nicht vorliegt. Wo Halluzinationen nicht vermieden werden können, ist ein entscheidungsersetzender Einsatz unzulässig. Diese Grenzen wahren derzeit nur regelbasierte Systeme. Vom Einsatz von Systemen maschinellen

Lernens ist im entscheidungsersetzenden Bereich derzeit eher abzuraten, um negative Auswirkungen auf die Akzeptanz der Rechtsschutzsuchenden und das Vertrauen in das Funktionieren des Rechtsstaates zu vermeiden; ihr Einsatz stünde zudem ohnehin derzeit mit der Gesetzesbindung nach Art. 20 Abs. 3 GG in Konflikt, deren Einhaltung systemseitig derzeit nicht sichergestellt werden kann. Erste und notwendige Handlungsoption ist vorliegend die Identifikation relevanter und möglicher Einsatzbereiche, z. B. Massenverfahren wie Fluggastrechteverfahren oder die Kostenfestsetzung nach §§ 103, 104 ZPO. Diese sollten durch interdisziplinär besetzte, Richterinnen und Richter einschließende, Arbeitsgruppen ermittelt und die Entscheidungskriterien in einem Entscheidungsbaum präzise herausgearbeitet werden; Systeme können zunächst für Teilbereiche entwickelt und sodann modular erweitert werden.
Strukturierte Abfrage- und Eingabemasken können zur Erhebung der relevanten Tatsachen genutzt werden, indes muss eine freie Äußerung möglich bleiben. Die Aussteuerung sollte bereits im Systemdesign angelegt sein, indem atypische, nicht systemseitig erfasste Konstellationen automatisiert ausgesteuert werden. Typische Sachverhalte sollten hingegen möglichst vordefiniert und systemseitig bearbeitet werden. Für die Aussteuerung auf Antrag sollten Maßnahmen zur Missbrauchsverhinderung vorgesehen werden. Hierbei kommt z. B. ein Substantiierungserfordernis in Betracht. Insgesamt sollte hierbei ein Ausgleich zwischen Praktikabilität, Zugänglichkeit und Missbrauchsverhinderung angestrebt werden. Der ausgesteuerte Fall sollte strukturiert aufbereitet an den Richter bzw. die Richterin übergeben werden. Verfahrensrechtlich kommt eine „0.-Instanz“ mit verbindlicher gerichtlicher Entscheidung oder eine dem Mahnverfahren ähnliche Ausgestaltung in Betracht, bei der die letztverbindliche Entscheidung auf Antrag dem Richter bzw. der Richterin obliegt. Empfohlen wird der entscheidungsersetzende Einsatz mittels regelbasierter Systeme im Sinne einer echten Instanz, um nachhaltige Effizienzgewinne zu bewirken.
Das „Ob“ und das „Wie“ des entscheidungsersetzenden Einsatzes bedürfen nach dem Demokratie- und dem Rechtsstaatsprinzip einer umfassenden Rechtsgrundlage, die die konkreten Einsatzbereiche und -modalitäten festlegt. Das Recht auf den gesetzlichen Richter verlangt zudem eine im Voraus feststehende Regelung, welche Richterinnen und Richter das ordnungsgemäße Funktionieren des Systems kontrollieren, es notfalls korrigieren und für Aussteuerungen zuständig sind. Datenschutzrechtlich ist eine Rechtsgrundlage (nach Art. 22 Abs. 2 lit. b) DSGVO) erforderlich, zur Überwindung des Verbots einer automatisierten Entscheidung nach Art. 22 Abs. 1 DSGVO. Es empfiehlt sich, diese in der ohnehin zu schaffenden gesetzlichen Grundlage mitzuregeln, anstatt, dass auf eine Einwilligung der Parteien gesetzt wird, deren taktische Verweigerung die Effizienzgewinne unterlaufen könnte. In der zu schaffenden Rechtsgrundlage sollten die wesentlichen Strukturentscheidungen ausdrücklich abgebildet werden. Sie sollte technologieoffen formuliert sein und zum Erlass einer Rechtsverordnung ermächtigen. Durch Anpassung von GVG und ZPO sind die für den entscheidungsersetzenden Einsatz notwendigen Regelungen zur Aussteuerung und Anpassung der sachlichen, örtlichen und funktionellen Zuständigkeit vorzusehen. Die sachliche Zuständigkeit sollte streitwertunabhängig und anhand der Materien zugewiesen werden. Eine bundesweit konzentrierte Zentralzuständigkeit bietet sich hierbei an. Auch örtlich kann der Einsatz entscheidungsersetzender Systeme zentral angesiedelt werden. Ausgesteuerte Verfahren könnten z. B. an das Gericht ausgesteuert werden, das ohne die Konzentration zuständig wäre und dort auch so verteilt werden, als wäre kein entscheidungsersetzendes System vorgesehen gewesen. Auch für Aufsicht, Kontrolle

und Korrektur des Systems sind Zuständigkeitsregelungen vorzusehen. Eine maßgebliche Möglichkeit besteht darin, das Vorhandensein einer interdisziplinären hohen Expertise für diese Tätigkeiten normativ festzulegen, wobei Optimierungspotentiale aus Aussteuerungs- und Rechtsmittelfällen für die Systemverbesserung genutzt werden sollten. Obgleich ein entscheidungsersetzender Einsatz derzeit politisch nicht erstrebt wird, sollten dennoch für die hinreichend einfach gelagerten Fälle die gesetzlichen Möglichkeiten für den entscheidungsersetzenden Einsatz geschaffen und damit die Entlastungsmöglichkeiten für die wichtige Ressource Justiz genutzt oder aber die Notwendigkeit des Fortbestehens derartiger Rechtsmaterien kritisch geprüft werden.
Der Rechtsschutz kann im Grundsatz durch das bestehende Rechtsmittelsystem gewährleistet werden. Gesondert gestaltete Rechtsbehelfe sind nicht zwingend vorzusehen. Die geltenden Berufungs- und Revisionsgründe erscheinen auch zur Überprüfung automatisiert getroffener Entscheidungen tauglich. Bei einer Ausgestaltung ähnlich dem Mahnverfahren kann auf Antrag eine Überführung des Vorgangs in ein streitiges Verfahren vor einer menschlichen Richterin oder einem menschlichen Richter vorgesehen werden. Eine echte „0.-Instanz“ erfordert die Schaffung einer neuen Berufungszuständigkeit, angesiedelt auf der Ebene der ersten Instanz. Die Zuständigkeit für Revisionen kann beim Bundesgerichtshof verbleiben oder mit einer Vorlagepflicht entsprechend § 121 Abs. 2 GVG den Oberlandesgerichten zugewiesen werden. Für die derzeit lediglich in Betracht kommenden hinreichend einfach gelagerten Fälle erscheint ein viergliedriger Instanzenzug eher unangemessen, sodass der Einsatz des entscheidungsersetzenden Systems auch als erste Instanz eines dreigliedrigen Instanzenzugs ausgestaltet werden kann. Anpassungsbedarf besteht bei der Statthaftigkeit: Verfahren über geringe Geldsummen, wie im Falle von Ausgleichsansprüchen nach der Fluggastrechte-VO, werden die Beschwerdegegenstands-Wertgrenze von 1000 Euro regelmäßig nicht überschreiten. Zur Wahrung insbesondere des Grundvertrauens in den Rechtsstaat sollten Berufungen gegen Entscheidungen, die durch entscheidungsersetzende informationstechnische Systeme getroffen wurden, stets zugelassen werden, statt sie dem Weg über eine problembehaftete Rechtsmittelzulassung auszusetzen. Ein Flagging-Tool im Justizportal kommt als zusätzlicher Rechtsbehelf optional in Betracht. Mit Blick auf §§ 41 ff. ZPO kann zudem eine gesetzliche Klarstellung erwogen werden.
Im entscheidungsersetzenden Einsatz sollte die Interaktion mit einem KI-System den Parteien gegenüber stets klar offengelegt werden. Soweit Parteien im entscheidungsersetzenden Einsatz ausschließlich mit dem System interagieren, kann der Verfahrensfortgang von einer Bestätigung der Kenntnisnahme der zu verschaffenden Informationen abhängig gemacht werden, jedoch bei der beklagten Partei an die Geltendmachung von Verteidigungsmitteln geknüpft. Um die im entscheidungsersetzenden Einsatz für entsprechend verfasste Entscheidungen einschlägige anbieterseitige Kennzeichnungspflicht sowie die im Regelfall einschlägige Offenlegungspflicht des Betreibers auszufüllen, können z. B. textliche und maschinenlesbare Hinweise darauf genutzt werden, dass die Entscheidung automatisiert durch ein KI-System generiert worden ist. Soweit der Zugang zu Gericht im entscheidungsersetzenden Szenario vornehmlich digital über eine Online-Plattform stattfindet, sind bei den Gerichten niedrigschwellige Unterstützungsmöglichkeiten für Personen ohne die notwendigen digitalen Kompetenzen oder Ressourcen einzurichten. Hierzu können bzw. sollten insbesondere physische Rechtsantragstellen geschaffen werden, die durch virtuelle Antragstellen ergänzt werden können. Anhand von fortgesetzter nutzerzentrierter sowie aus Nutzererfahrungen lernender

Maßnahmen sollten Nutzerbedürfnisse in einem iterativen Entwicklungsprozess stets besser umgesetzt werden. Der digitale Zugang sollte umfassend barrierefrei ausgestaltet werden. Auch sollte insbesondere eine Absenkung der Gerichtskosten erwogen werden.

e. Gesamtschau und Ausblick

Die im ersten Arbeitspaket (Milestone 1) des vorliegenden Forschungsprojekts erforschten und im vorliegenden Arbeitspaket (Milestone 2 & 3) aggregierten rechtlichen roten Linien geben den Rahmen vor, innerhalb dessen der Einsatz informationstechnischer Systeme bei der richterlichen Entscheidungsfindung rechtlich möglich ist. Dieser Rahmen kann durch die in dem vorliegenden Arbeitspaket (Milestone 2 & 3) aufgezeigten Handlungsoptionen und sollte durch die in diesem aufgezeigten Handlungsempfehlungen ausgefüllt werden. Zwar birgt die Nutzung informationstechnischer Systeme bei der richterlichen Entscheidungsfindung Risiken, jedoch verursacht auch die Nichtnutzung Kosten, da Potentiale zur Einsparung von Zeit und Kosten ungenutzt und damit auch Personalressourcen weiterhin an den entsprechenden Punkten gebunden bleiben.
Die gegebenen Handlungsempfehlungen orientieren sich nicht allein an der rechtlichen Zulässigkeit. In die Auswahl der jeweiligen Handlungsempfehlungen sind wiederholt Aspekte der Akzeptanz, des Vertrauens und der Entlastungswirkung eingeflossen. Dies zeigt sich beispielhaft daran, dass bewusst eingebaute Fehler im Komplex der Ausgabenkontrolle trotz Zulässigkeit aus Akzeptanzgründen verworfen oder Begründungslasten unter einen Testvorbehalt gestellt wurden, eine Empfehlung zur weitestgehenden Zulassung und Stärkung der datenschutzrechtlichen Betroffenenrechte ausgesprochen wurde und bei der Überwindung des Verbots automatisierter Entscheidungen im entscheidungsersetzenden Einsatz eine gesetzliche Grundlage gegenüber der Einwilligungslösung präferiert wurde.
Als übergreifende Topoi kehren im vorliegenden Berichtsteil des vorliegenden Arbeitspakets (Milestone 2 & 3) insbesondere die Ansiedelung der Entwicklung und des Betriebs des informationstechnischen Systems bei der Judikative als „Königsweg“, die ganzheitliche Adressierung des gesamten Lebenszyklus der informationstechnischen Systeme sowie die Präferenz für Konzentration und Bund/Länder-übergreifende Abstimmungen wieder. Auch wird die Implementierung der Anforderungen der KI-VO über ihren zwingenden Anwendungsbereich hinaus wiederholt herangezogen.
Zahlreiche Befunde der Arbeitspakete des vorliegenden Forschungsprojekts sind ausdrücklich an den Stand gegenwärtiger Gegebenheiten geknüpft. Entsprechend sollte der Komplex der Künstlichen Intelligenz und richterlichen Entscheidungsfindung auch weiterhin fortlaufend eng begleitet werden. Für die Umsetzung des Einsatzes informationstechnischer Systeme bei der richterlichen Entscheidungsfindung sollten zunächst Fallgestaltungen, in denen eine Entscheidung auf eine Vielzahl von Fällen angewendet werden kann, prioritär anvisiert werden. Eine schrittweise Einführung kann hierbei eine früh einsetzende und sukzessiv ansteigende Entlastungswirkung bewirken. Die Funktionsfähigkeit des Rechtsstaats ist von kaum zu überschätzender Bedeutung. Informationstechnische Systeme bei der richterlichen Entscheidungsfindung können ihren Beitrag hierzu leisten. Dies sollte genutzt werden.

III. Zusammenfassende Schlussfolgerungen

Die Ergebnisse des Forschungsprojekts „Künstliche Intelligenz und richterliche Entscheidungsfindung“ verdeutlichen, dass der Einsatz Künstlicher Intelligenz in der Justiz erhebliche Potenziale eröffnet, zugleich aber besondere Anforderungen an die technische Gestaltung, die institutionelle Einbettung und die rechtliche Absicherung entsprechender Systeme stellt. Die technisch-ethische und die rechtliche Untersuchung gelangen dabei aus unterschiedlichen Ausgangsperspektiven zu einer gemeinsamen zentralen Erkenntnis: **Die Vertretbarkeit und Zulässigkeit des KI-Einsatzes können nicht abstrakt anhand der verwendeten Technologie beurteilt werden, sondern bestimmen sich maßgeblich nach dem funktionalen Charakter der Tätigkeit, seinem Automatisierungsgrad und der Komplexität des Anwendungsfeldes.** Auch die Bedingungen der institutionellen Einbettung des Systems agieren als hierbei wesentlicher Faktor. Ein verantwortungsvoller KI-Einsatz erfordert daher eine differenzierte und anwendungsbezogene Betrachtung, die technische Leistungsfähigkeit, menschliche Entscheidungsprozesse, organisatorische Rahmenbedingungen und rechtsstaatliche Anforderungen gleichermaßen berücksichtigt.

Ausgangspunkt beider Untersuchungsperspektiven ist die Erkenntnis, dass weder eine pauschale Ablehnung noch eine unkritische Befürwortung des KI-Einsatzes sachgerecht erscheint. Die Nutzung informationstechnischer Systeme kann zur Entlastung der Gerichte, zur Beschleunigung von Verfahren und zur Verbesserung der Informationsverarbeitung beitragen. Zugleich können neue Fehlerquellen, Abhängigkeiten und strukturelle Verzerrungen entstehen. Insbesondere die Möglichkeit, dass sich fehlerhafte oder diskriminierende Systemergebnisse über zahlreiche Verfahren hinweg reproduzieren, erfordert eine Betrachtung über die einzelne gerichtliche Entscheidung hinaus. Ebenso sind die Opportunitätskosten der Nichtnutzung zu berücksichtigen, soweit sich entsprechende Effizienzpotenziale empirisch nachweisen lassen. Die Bewertung des KI-Einsatzes muss deshalb die Risiken seiner Nutzung und die Risiken des bestehenden Zustands gleichermaßen einbeziehen, ohne dabei rechtsstaatlich unverzichtbare Anforderungen einer bloßen Effizienzabwägung zugänglich zu machen.

Eine wesentliche gemeinsame Schlussfolgerung betrifft die **zentrale Bedeutung wirksamer menschlicher Aufsicht und Kontrolle**. Aus technisch-ethischer Sicht genügt es nicht, menschlichen Entscheidungsträgern lediglich formal die Letztentscheidung oder eine Korrekturmöglichkeit zuzuweisen. Effektive Kontrolle setzt tatsächliche Überprüfungs- und Eingriffsmöglichkeiten, ausreichende Kenntnisse über Funktionsweise und Grenzen der Systeme sowie geeignete institutionelle Rahmenbedingungen voraus. Dies korrespondiert rechtlich insbesondere mit den Anforderungen der richterlichen Unabhängigkeit und der Gesetzesbindung sowie zudem mit der Notwendigkeit, eine verantwortliche Kontrolle der eingesetzten Systeme sicherzustellen. Richterinnen und Richtern müssen hierzu effektive und substantielle Kontroll- und Korrekturmöglichkeiten zur Verfügung gestellt werden. Auch entsprechende Aus- und Fortbildungsmaßnahmen rechnen hierzu, ebenso wie prozedurale Gegenimpulse gegen Technikübervertrauen. Es besteht somit die Notwendigkeit ineinandergreifender und abgestimmter technischer, organisatorischer und rechtlicher Schutzmaßnahmen.

Die Untersuchungen zeigen ferner, dass **die Gestaltung und Governance von KI-Systemen ein wesentlicher Faktor rechtsstaatlicher Entscheidungsfindung** ist. Technische Systeme beeinflussen nicht nur durch ihre Ergebnisse, sondern bereits durch die (Vor-)Strukturierung von Informationen, Benutzeroberflächen und Handlungsmöglichkeiten die richterliche Entscheidungsfindung. Ausgaben informationstechnischer Systeme, insbesondere algorithmische Empfehlungen können psychologische (Anker)Effekte hervorrufen bzw. mit der Problematik dieser verbunden sein und faktische Entscheidungsvorgaben erzeugen. Dem ist durch transparente, herkunftssensitive Benutzeroberflächen, angemessene Kontrollmechanismen und eine institutionell geschützte Entscheidungsfreiheit zu begegnen. Zugleich ist eine wirksame Beteiligung der Richterschaft an Beschaffung, Betrieb, Überwachung und – je nach Funktion – auch der Entwicklung der Systeme erforderlich. Die konkrete Ausgestaltung und Intensität hängt hierbei von dem Grad des funktionalen Charakters der Tätigkeit, dem Grads der Automatisierung und der Komplexität des Anwendungsfeldes ab. Je stärker das System sachlich-inhaltliche Aufgaben wahrnimmt, je höher der Grad der Automatisierung ist und je komplexer das Anwendungsfeld gestaltet ist, desto höher gestalten sich die notwendigen Anforderungen. (Auch) Die Ansiedelung von Entwicklung und Betrieb innerhalb der Judikative erscheint insoweit als besonders geeigneter Weg, der zudem technologische Souveränität und institutionelle Unabhängigkeit zu gewährleisten verspricht.

Hiermit verbunden ist die Erkenntnis, dass die Verantwortung für einen rechtmäßigen und technisch-ethisch vertretbaren KI-Einsatz nicht allein bei den einzelnen Anwenderinnen und Anwendern liegen kann. Vielmehr bedarf es einer **funktionalen Verantwortungsverteilung** zwischen Richterschaft, Systementwicklung, Betrieb und Organisation, die auch rechtlich in unterschiedlichen Rahmenregelungen bereits vorgezeichnet ist. Systemische Risiken lassen sich ferner regelmäßig nicht durch punktuelle Kontrollen einzelner Ergebnisse beherrschen. Erforderlich sind vielmehr kontinuierliche Risikoanalysen, Qualitätssicherungsmechanismen, diskriminierungssensible Datenverarbeitung, wirksame IT-Sicherheit sowie institutionalisierte Monitoring- und Feedbackprozesse. Unabhängige Prüfungen und interdisziplinäre Evaluationsstrukturen sollten diese Maßnahmen ergänzen. Der technisch-ethisch entwickelte Evaluations- und Reflexionsrahmen und die rechtlich empfohlenen Governance- und Kontrollstrukturen bilden insoweit komplementäre Instrumente einer umfassenden Risikosteuerung.

Für die konkrete Implementierung folgt aus beiden Untersuchungsteilen die Notwendigkeit eines **abgestuften, funktions- und risikoorientierten Vorgehens**. Entscheidungsunterstützende Systeme können insbesondere dort erhebliche Vorteile bieten, wo umfangreiche Informationen strukturiert, aufbereitet und für die richterliche Entscheidung zugänglich gemacht werden müssen. Voraussetzung bleibt, dass Richterinnen und Richter die Ergebnisse eigenständig überprüfen und korrigieren können und ihre Entscheidungsfreiheit gewahrt bleibt. Ein entscheidungsersetzender Einsatz unterliegt demgegenüber deutlich strengeren Anforderungen und kommt nach gegenwärtigem Erkenntnisstand nur für hinreichend einfach gelagerte, weitgehend determinierte Fallkonstellationen in Betracht. Die für diese Einsatzart notwendigen Anforderungen können derzeit lediglich regelbasierte Systeme einhalten. Unverzichtbar sind für einen entscheidungsersetzenden Einsatz insbesondere eine umfassende gesetzliche Grundlage, eindeutig geregelte Zuständigkeiten, wirksame Aussteuerungsmöglichkeiten zu menschlichen Richterinnen und Richtern (, wenn sich

herausstellt, dass kein hinreichend einfach gelagerter Fall in diesem Sinne vorliegt) sowie ein effektiver gerichtlicher Rechtsschutz.

Eine weitere übergreifende Schlussfolgerung betrifft die **Notwendigkeit einer dauerhaft lern- und anpassungsfähigen Justizorganisation**. Die dynamische Weiterentwicklung Künstlicher Intelligenz führt dazu, dass technische Möglichkeiten, Risikoprofile und tatsächliche Nutzungskontexte einem fortlaufenden Wandel unterliegen. Statische Bewertungen und einmalige Zulassungsentscheidungen können diesen Entwicklungen nicht hinreichend Rechnung tragen. Erforderlich sind deshalb kontinuierliche Evaluationen, regelmäßige Anpassungen technischer und organisatorischer Schutzmechanismen sowie der Aufbau eigenständiger justizieller Kompetenzen. Besonders die Entwicklung zunehmend autonom handelnder KI-Systeme verdeutlicht, dass bestehende Konzepte menschlicher Aufsicht künftig vor neue Herausforderungen gestellt werden können. Die Justiz muss daher in die Lage versetzt werden, technologische Entwicklungen selbständig und unabhängig anhand stabiler rechtsstaatlicher Maßstäbe zu beurteilen.

Die technisch-ethischen und rechtlichen Handlungsempfehlungen zeigen Wege auf, wie die identifizierten Anforderungen für den Einsatz von Künstlicher Intelligenz bei der richterlichen Entscheidungsfindung sachgerecht und praxistauglich verwirklicht werden können. Entscheidend wird letztlich sein, die identifizierten Maßnahmen in tragfähige rechtliche, technische und institutionelle Strukturen zu überführen und deren Wirksamkeit fortlaufend zu überprüfen. Unter diesen Voraussetzungen kann Künstliche Intelligenz einen wichtigen Beitrag zur langfristigen Leistungs- und Funktionsfähigkeit der Justiz leisten, ohne die grundlegenden Prinzipien rechtsstaatlicher richterlicher Entscheidungsfindung zu gefährden. Die Funktionsfähigkeit dieses Rechtsstaats ist von kaum zu überschätzender Bedeutung.

LITERATURVERZEICHNIS

I. Erstes Arbeitspaket (Milestone 1) des Forschungsprojekts „Künstliche Intelligenz und richterliche Entscheidungsfindung“

Gold, Valentin/Reuß, Philipp M./del Rio Hecklau, Jana K./Gätjen, Franz K./Kohler, Joshua A. K./Kohler, Katja I./Kolbe, Moritz/Thamm, Thomas W./Thomsen, Henrike J., FORSCHUNGSPROJEKT Künstliche Intelligenz und richterliche Entscheidungsfindung, Milestone 1: Abschlussbericht, Göttingen 2025, abrufbar unter https://reusz.eu/upload/maki_m1_abschlussbericht.pdf (Abruf am 24.10.2026, 22:14 Uhr) (zit. als Gold/Reuß et al., Forschungsprojekt KI und richterliche Entscheidungsfindung, 2025, S. Seitenzahl).

II. Monographien

Adolphsen, Jens, Zivilprozessrecht, 9. Aufl., Baden-Baden 2025.

Fenton, Norman/Neil, Martin, Risk Assessment and Decision Analysis with Bayesian Networks, 1. Aufl., Boca Raton 2012.

Fischer, John Martin/Ravizza, Mark, Responsibility and Control – A Theory of Moral Responsibility, 1. Aufl., Cambridge 1998.

Förster, Christian/Straburzynski, Julia, Die KI-Verordnung in der Praxis – Rechtliche Grundlagen und Pflichten bei der Anwendung von KI im Unternehmen, 1. Aufl., München 2025.

Graichen, Nikolas, Die Automatisierung der Justiz, 1. Aufl., Baden-Baden 2022.

Kahneman, Daniel, Thinking, Fast and Slow, 1. Aufl., New York 2011.

Kernchen, Tanja, Automatisierung und richterliche Rechtsanwendung – Rechtliche Anforderungen und technische Möglichkeiten der Entwicklung eines intelligenten Systems für richterliche Rechtsanwendung, 1. Aufl., Baden-Baden 2025.

Kindhäuser, Urs/Schumann, Kay H., Strafprozessrecht, 8. Aufl., Baden-Baden 2025.

Köhler, Moritz, Richterliche Verantwortung und Künstliche Intelligenz – Zur Abgrenzung zwischen unzulässiger Ersetzung und zulässiger Unterstützung des menschlichen Richters, 1. Aufl., Baden-Baden 2026.

Korb, Kevin B./Nicholson, Ann E., Bayesian Artificial Intelligence, 2. Aufl., Boca Raton 2010.

Martini, Mario, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz, 1. Aufl., Berlin, Heidelberg 2019.

Martini, Mario/Botta, Jonas/Nink, David/Kolain, Michael, Automatisch erlaubt? Fünf Anwendungsfälle algorithmischer Systeme auf dem juristischen Prüfstand, Bertelsmann-Stiftung (Hrsg.), Gütersloh 2020.

Nink, David, Justiz und Algorithmen – Über die Schwächen menschlicher Entscheidungsfindung und die Möglichkeiten neuer Technologien in der Rechtsprechung, 1. Aufl., Berlin 2021.

Oberheim, Rainer, Zivilprozessrecht für Referendare, 15. Aufl., München 2023.

Ostendorf, Heribert/Brüning, Janique, Strafprozessrecht, 5. Aufl., Baden-Baden 2024.

Quack, Lukas, Automation zivilrechtlicher Entscheidungsfindung, 1. Aufl., Baden-Baden 2025.

Reason, James, Human Error, 1. Aufl., Cambridge 1990.

Rollberg, Christoph, Algorithmen in der Justiz – Rechtsfragen zum Einsatz von Legal Tech im Zivilprozess, 1. Aufl., Baden-Baden 2020.

Rosenberg, Leo/Schwab, Karl Heinz/Gottwald, Peter, Zivilprozessrecht, 18. Aufl., München 2018.

Wolff, Lydia, Algorithmen als Richter – Verfassungsrechtliche Grenzen entscheidungstreffender Rechtsgeneratoren in der Rechtsprechung, 1. Aufl., Trier 2022.

III. Sammelwerke

Bräutigam, Peter (Hrsg.), IT-Outsourcing und Cloud-Computing – Eine Darstellung aus rechtlicher, technischer, wirtschaftlicher und vertraglicher Sicht, 4. Aufl., Berlin 2019 (zit. als *Bearbeiter*, in: Bräutigam (Hrsg.), IT-Outsourcing und Cloud-Computing, Teil *Teil* Rn. *Randnummer*).

Bystranowski, Piotr/Janik, Bartosz/Próchnicki, Maciej (Hrsg.), Judicial Decision-Making – Integrating Empirical and Theoretical Perspectives, Economic Analysis of Law in European Legal Scholarship, Bd. 14, 1. Aufl., Cham 2022; darin: *Próchnicki, Maciej/Janik, Bartosz/Bystranowski, Piotr*, Debiasing Numerical Verdicts and Judicial Discretion: Reflections on Mitigating the Anchoring Effect in Judicial Decision-Making, S. 79–103 (zit. als *Próchnicki/Janik/Bystranowski*, Debiasing Numerical Verdicts and Judicial Discretion: Reflections on Mitigating the Anchoring Effect in Judicial Decision-Making, in: Bystranowski/Janik/Próchnicki (Hrsg.), Judicial Decision-Making – Integrating Empirical and Theoretical Perspectives, 1. Aufl., Cham 2022, S. *Seitenzahl*).

Chibanguza, Kuuya J./Kuß, Christian/Steege, Hans (Hrsg.), Künstliche Intelligenz – Recht und Praxis automatisierter und autonomer Systeme, 2. Aufl., Baden-Baden 2026 (zit. als Chibanguza/Kuß/Steege KI-Hdb/*Bearbeiter*, 2. Aufl. 2026, *Teil*: § *Paragraph Römischer Großbuchstabe*. Rn. *Randnummer*).

Ebers, Martin/Heinze, Christian/Steinrötter, Björn (Hrsg.), Künstliche Intelligenz und Robotik – Rechtshandbuch, 2. Aufl., München 2026 (zit. als Ebers/Heinze/Steinrötter KI/*Bearbeiter*, 2. Aufl. 2026, § *Paragraph* Rn. *Randnummer*).

Hermann, Dieter/Pöge, Andreas (Hrsg.), Kriminalsoziologie – Handbuch für Wissenschaft und Praxis, 1. Aufl., Baden-Baden 2018; darin: *Grundies, Volker*, Regionale Unterschiede in der gerichtlichen Sanktionspraxis in der Bundesrepublik Deutschland. Eine empirische Analyse, S. 295–316 (zit. als *Grundies*, Regionale Unterschiede in der gerichtlichen Sanktionspraxis in der Bundesrepublik Deutschland. Eine empirische Analyse, in: Hermann/Pöge (Hrsg.), Kriminalsoziologie – Handbuch für Wissenschaft und Praxis, 1. Aufl., Baden-Baden 2018, S. *Seitenzahl*).

Hilgendorf, Eric/Roth-Isigkeit, David (Hrsg.), Die neue Verordnung der EU zur Künstlichen Intelligenz – Rechtsfragen und Compliance, 2. Aufl., München 2025 (zit. als Hilgendorf/Roth-Isigkeit KI-VO/*Bearbeiter*, 2. Aufl. 2025, § *Paragraph* Rn. *Randnummer*).

Hornung, Gerrit/Schallbruch, Martin (Hrsg.), IT-Sicherheitsrecht – Praxishandbuch, 2. Aufl., Baden-Baden 2025 (zit. als Hornung/Schallbruch/*Bearbeiter*, IT-Sicherheitsrecht, 2. Aufl. 2024, § *Paragraph arabischer Großbuchstabe*. *römische Ziffer*. *Arabische Ziffer*. Rn. *Randnummer*).

Kaulartz, Markus/Braegelmann, Tom (Hrsg.), Rechtshandbuch Artificial Intelligence und Machine Learning, 1. Aufl., München 2020 (zit. als Kaulartz/Braegelmann AI und Machine Learning-HdB/*Bearbeiter*, 1. Aufl. 2020, Kap. *Kapitel* Rn. *Randnummer*.).

Marly, Jochen/Bomhard, David/Schreiber, Kristina (Hrsg.), Praxishandbuch Softwarerecht – Rechtsschutz und Vertragsgestaltung, 8. Aufl., München 2024 (zit. als Marly SoftwareR-HdB/*Bearbeiter*, 8. Aufl. 2024, § *Paragraph* Rn. *Randnummer*).

Quiñonero-Candela, Joaquin/Sugiyama, Masashi/Schwaighofer, Anton/Lawrence, Neil D. (Hrsg.), Dataset Shift in Machine Learning, Neural Information Processing Series, 1. Aufl., Cambridge (Mass.) 2008, abrufbar unter https://direct.mit.edu/books/edited-volume/3841/Dataset-Shift-in-Machine-Learning (Abruf am 10.09.2026, 16:22 Uhr).

Sousa Antunes, Henrique/Freitas, Pedro Miguel/Oliveira, Arlindo L./Martins Pereira, Clara/Vaz de Sequeira, Elsa/Barreto Xavier, Luís (Hrsg.), Multidisciplinary Perspectives on Artificial Intelligence and the Law, Law, Governance and Technology Series, Bd. 58, 1. Aufl., Cham 2024; darin: *Gravett, Willem H.*, Judicial Decision-Making in the Age of Artificial Intelligence, S. 281–297 (zit. als *Gravett*, Judicial Decision-Making in the Age of Artificial Intelligence, in: Sousa Antunes/Freitas/Oliveira/Martins Pereira/Vaz de Sequeira/Barreto Xavier (Hrsg.), Multidisciplinary Perspectives on Artificial Intelligence and the Law, 1. Aufl., Cham 2024, S. *Seitenzahl*).

IV. Kommentare

Bomhard, David/Pieper, Fritz-Ulli/Wende, Susanne (Hrsg.), Kommentar KI-VO – Verordnung über künstliche Intelligenz, 1. Aufl., Frankfurt a. M. 2025. (zit. als: Bomhard/Pieper/Wende/*Bearbeiter*, 1. Aufl. 2025, KI-VO Art. *Artikel* Rn. *Randnummer*)

Dreier, Horst (Hrsg.), Grundgesetz Kommentar, Bd. 2: Artikel 20–82, 3. Aufl., Tübingen 2015. (zit. als Dreier/*Bearbeiter*, 3. Aufl. 2015, GG Art. *Artikel* Artikel *Artikel* Rn. *Randnummer*)

Dürig, Günter/Herzog, Roman/Scholz, Rupert (Begr./Hrsg.), Grundgesetz, Kommentar, Loseblatt, 111. EL, München Juni 2026. (zit. als Dürig/Herzog/Scholz/*Bearbeiter*, 111. EL Juni 2026, GG Art. *Artikel* Rn. *Randnummer*)

Ehmann, Eugen/Selmayr, Martin (Hrsg.), DS-GVO – Datenschutz-Grundverordnung – Kommentar, 3. Aufl., München 2024. (zit. als Ehmann/Selmayr/*Bearbeiter*, 3. Aufl. 2024, DS-GVO Art. *Artikel* Rn. *Randnummer*)

Epping, Volker/Hillgruber, Christian (Hrsg.), Beck’scher Online-Kommentar Grundgesetz, 66. Edition, München 2026. (zit. als BeckOK GG/*Bearbeiter*, 66. Ed. 15.8.2026, GG Art. *Artikel* Rn. *Randnummer*)

Graf, Jürgen Peter (Hrsg.), Beck’scher Online-Kommentar GVG, 32. Edition, München 2026. (zit. als BeckOK GVG/*Bearbeiter*, 32. Ed. 15.8.2026, GVG § *Paragraph* Rn. *Randnumer*)

Heckmann, Dirk/Rost, Maria Christina (Hrsg.), Datenschutz-Grundverordnung VO (EU) 2016/679, Bundesdatenschutzgesetz: DS-GVO / BDSG, 4. Aufl., München 2026. (zit. als Gola/Heckmann/Rost/*Bearbeiter*, 4. Aufl. 2026, DS-GVO Art. *Artikel* Rn. *Randnummer*)

Hilgendorf, Eric/Härtlein, Johannes (Hrsg.), KI-VO – Verordnung über künstliche Intelligenz, 1. Aufl., Baden-Baden 2025. (zit. als HK-KI-VO/Hilgendorf/Härtlein/*Bearbeiter*, 1. Aufl. 2025, KI-VO Art. *Artikel* Rn. *Randnummer*)

Jarass, Hans D./Kment, Martin, Grundgesetz für die Bundesrepublik Deutschland – Kommentar, 19. Aufl., München 2026. (zit. als Jarass/Pieroth/*Bearbeiter*, 19. Aufl. 2026, GG Art. *Artikel* Rn. *Randnummer*)

Krüger, Wolfgang/Rauscher, Thomas (Hrsg.), Münchener Kommentar zur Zivilprozessordnung mit Gerichtsverfassungsgesetz und Nebengesetzen, Bd. 1: §§ 1–354, 7. Aufl., München 2025. (zit. als MüKoZPO/*Bearbeiter*, 7. Aufl. 2025, ZPO § Paragraph Rn. *Randnummer*)

Krüger, Wolfgang/Rauscher, Thomas (Hrsg.), Münchener Kommentar zur Zivilprozessordnung mit Gerichtsverfassungsgesetz und Nebengesetzen, Bd. 2: §§ 355–945b, 7. Aufl., München 2025. (zit. als MüKoZPO/*Bearbeiter*, 7. Aufl. 2025, ZPO § Paragraph Rn. *Randnummer*)

Krüger, Wolfgang/Rauscher, Thomas (Hrsg.), Münchener Kommentar zur Zivilprozessordnung mit Gerichtsverfassungsgesetz und Nebengesetzen, Bd. 3: §§ 964–1120, 6. Aufl., München 2022. (zit. als MüKoZPO/*Bearbeiter*, 6. Aufl. 2022, GVG § *Paragraph* Rn. *Randnummer*)

Kühling, Jürgen/Buchner, Benedikt (Hrsg.), Datenschutz-Grundverordnung/BDSG – Kommentar, 4. Aufl., München 2024. (zit. als Kühling/Buchner/*Bearbeiter*, 4. Aufl. 2024, DS-GVO Art. *Artikel* Rn. *Randnummer*)

Martini, Mario/Wendehorst, Christiane (Hrsg.), KI-VO – Verordnung über künstliche Intelligenz, Kommentar, 1. Aufl., München 2024. (zit. als Martini/Wendehorst/*Bearbeiter*, 1. Aufl. 2024, KI-VO Art. *Artikel* Rn. *Randnummer*)

Martini, Mario/Wendehorst, Christiane (Hrsg.), KI-VO – Verordnung über künstliche Intelligenz, Kommentar, 2. Aufl., München 2026. (zit. als Martini/Wendehorst/*Bearbeiter*, 2. Aufl. 2026, KI-VO Art. *Artikel* Rn. *Randnummer*)

Musielak, Hans-Joachim/Voit, Wolfgang, Zivilprozessordnung mit Gerichtsverfassungsgesetz, Kommentar, 23. Aufl., München 2026. (zit. als Musielak/Voit/*Bearbeiter*, 23. Aufl. 2026, *Gesetz* § *Paragraph* Rn. *Randnummer*)

Ory, Stephan/Weth, Stephan (Hrsg.), juris PraxisKommentar Elektronischer Rechtsverkehr – Band 1: Allgemeiner Teil, 3. Aufl., Saarbrücken 2025. (zit. als: *Bearbeiter* in: Ory/Weth, jurisPK-ERV Band 1, 3. Aufl., Kapitel *Kapitel* (Stand: *Datum des Standes*), Rn. *Randnummer)*

Paal, Boris P./Pauly, Daniel A. (Hrsg.), Datenschutz-Grundverordnung, Bundesdatenschutzgesetz, 4. Aufl., München 2026. (zit. als Paal/Pauly/*Bearbeiter*, 4. Aufl. 2026, *Gesetz/Verordnung* Art. *Artikel* Rn. *Randnummer*)

Reusch, Philipp/Chibanguza, Kuuya (Hrsg.), KI-VO – Verordnung über Künstliche Intelligenz, 1. Aufl., München 2026. (zit. als Reusch/Chibanguza/Bearbeiter, 1. Aufl. 2026, KI-VO Art. *Artikel* Rn. *Randnummer*)

Saenger, Ingo (Hrsg.), Zivilprozessordnung Familienverfahren | Gerichtsverfassung | Europäisches Verfahrensrecht, 10. Aufl., Baden-Baden 2023. (zit. als Saenger/*Bearbeiter*, 10. Aufl. 2023, *Gesetz* § *Paragraph* Rn. *Randnummer*)

Schefzig, Jens/Kilian, Robert (Hrsg.), Beck’scher Online-Kommentar KI-Recht, 7. Edition, München 2026. (zit. als BeckOK KI-Recht/*Bearbeiter*, 7. Ed. *Stand-Datum*, KI-VO Art. *Artikel* Rn. *Randnummer*)

Simitis, Spiros/Hornung, Gerrit/Spiecker gen. Döhmann, Indra (Hrsg.), Datenschutzrecht – DS-GVO/BDSG, 2. Aufl., Baden-Baden 2025. (zit. als Simitis/Hornung/Spiecker gen. Döhmann/*Bearbeiter*, 2. Aufl. 2025, *Gesetz/Verordnung* § *Paragraph* Rn. *Randnummer*)

Skauradszun, Dominik (Hrsg.), Verbraucherrechtedurchsetzungsgesetz (VDuG) – Kommentar, 1. Aufl., München 2024. (zit. als Skauradszun/*Bearbeiter*, 1. Aufl. 2024, VDuG § *Paragraph* Rn. *Randnummer*)

Stein, Friedrich (Begr.), Bork, Reinhard/Roth, Herbert (Hrsg.), Kommentar zur Zivilprozessordnung, Bd. 1: §§ 1–77, 24. Aufl., Tübingen 2024. (zit. als Stein/*Bearbeiter*, 24. Aufl. 2024, ZPO § *Paragraph* Rn. *Randnummer*)

Stein, Friedrich (Begr.), Bork, Reinhard/Roth, Herbert (Hrsg.), Kommentar zur Zivilprozessordnung, Bd. 3: §§ 148–270, 24. Aufl., München 2024. (zit. als Stein/*Bearbeiter*, 24. Aufl. 2024, ZPO § *Paragraph* Rn. *Randnummer*)

Taeger, Jürgen/Gabel, Detlev (Hrsg.), Kommentar DSGVO – BDSG – TDDDG, 5. Aufl., Frankfurt a. M. 2026. (zit. als Taeger/Gabel/*Bearbeiter*, 5. Aufl. 2026, DS-GVO Art. *Artikel* Rn. *Randnummer*)

Thomas, Heinz/Putzo, Hans (Begr.), Zivilprozessordnung, FamFG Verfahren in Familiensachen, EGZPO, GVG, EGGVG, VDuG, EU-Zivilverfahrensrecht, 47. Aufl., München 2026. (zit. als Thomas/Putzo/*Bearbeiter* ZPO, 47. Aufl. 2026, § *Paragraph* Rn. *Randnummer*)

Vorwerk, Volkert/Wolf, Christian (Hrsg.), Beck’scher Online-Kommentar ZPO, 61. Edition, München 2026. (zit. als BeckOK ZPO/*Bearbeiter*, 61. Ed. 1.7.2026, ZPO § Paragraph Rn. *Randnummer*)

Wolff, Heinrich Amadeus/Brink, Stefan/Ungern-Sternberg, Antje von (Hrsg.), Beck’scher Online-Kommentar Datenschutzrecht – DS-GVO, DA, DGA, BDSG. Datenschutz und Datennutzung, 57. Edition, München 2026. (zit. als BeckOK DatenschutzR/*Bearbeiter*, 57. Ed. 1.8.2026, *Gesetz/Verordnung* Art./§ *Artikel/Paragraph* Rn. *Randnummer*)

V. Tagungsbände

Bernzen, Anna/Heinze, Christian/Steinrötter, Björn (Hrsg.), Digitalregulierung in der Bewährung – DSRI-Tagungsband Herbstakademie 2025, München 2025; darin: *Bunes, Florian*, Schlau macht sicher – KI-Kompetenz als Compliance-Faktor, S. 371-379. (zit. als *Bunes*, in: Bernzen/Heinze/Steinrötter, DSRI Herbstakademie 2025, 1. Aufl. 2025, S. *Seitenzahl*).

Räckers, Michael/Halsbenning, Sebastian/Rätz, Detlef/Richter, David/Schweighofer, Erich (Hrsg.), Digitalisierung von Staat und Verwaltung, Lecture Notes in Informatics (LNI), Gesellschaft für Informatik e.V., Bonn 2019; darin: *Süptitz, Thomas/Eymann, Torsten*, IT-Outsourcing für die Justiz: Ein Eingriff in die richterliche Unabhängigkeit?, S. 15-26 (zit. als *Süptitz/Eymann*, IT-Outsourcing für die Justiz: Ein Eingriff in die

richterliche Unabhängigkeit?, in: Räckers/Halsbenning/Rätz/Richter/Schweighofer (Hrsg.), Digitalisierung von Staat und Verwaltung, Lecture Notes in Informatics (LNI), Bonn 2019, S. *Seitenzahl*).

Reuß, Philipp M./Laß, Jessica (Hrsg.), Göttinger Kolloquien zur Digitalisierung des Zivilverfahrensrechts, Bd. 3: Kolloquien im Sommersemester 2022 | Kolloquien im Wintersemester 2022/2023, Göttingen 2024; darin: *Spoenle, Jan*, OLGA – der KI-Assistent für Dieselverfahren, S. 169–176 (zit. als *Spoenle*, OLGA – der KI-Assistent für Dieselverfahren, in: Reuß/Laß (Hrsg.), Göttinger Kolloquien zur Digitalisierung des Zivilverfahrensrechts Kolloquien im Sommersemester 2022 | Kolloquien im Wintersemester 2022/2023, Göttingen 2024, S. *Seitenzahl*).

Reuß, Philipp M./Laß, Jessica (Hrsg.), Göttinger Kolloquien zur Digitalisierung des Zivilverfahrensrechts, Bd. 5: Kolloquien im Sommersemester 2024 | Kolloquien im Wintersemester 2024/2025, Göttingen 2026; darin: *Langer, Markus/Ruschemeier, Hannah/Kolbe, Moritz/Laux, Johann*, „Die KI wird’s schon richten“ – Eine psychologische Perspektive auf Automation Bias und richterliche Unabhängigkeit, S. 79-99 (zit. als *Langer/Ruschemeier/Kolbe/Laux*, „Die KI wird’s schon richten“ – Eine psychologische Perspektive auf Automation Bias und richterliche Unabhängigkeit, in: Reuß/Laß (Hrsg.), Göttinger Kolloquien zur Digitalisierung des Zivilverfahrensrechts Kolloquien im Sommersemester 2024 | Kolloquien im Wintersemester 2024/2025, Göttingen 2026, S. *Seitenzahl*).

Reuß, Philipp M./Windau, Benedikt (Hrsg.), Göttinger Kolloquien zur Digitalisierung des Zivilverfahrensrechts, Bd. 1: Kolloquien im Sommersemester 2021, Göttingen 2022; darin: *Dörr, Sina*, Elektronischer Rechtsverkehr und Bürgerzugänge, S. 113–122 (zit. als *Dörr*, Elektronischer Rechtsverkehr und Bürgerzugänge, in: Reuß/Windau (Hrsg.), Göttinger Kolloquien zur Digitalisierung des Zivilverfahrensrechts, Kolloquien im Sommersemester 2021, Göttingen 2022, S. *Seitenzahl*).

Schweighofer, Erich/Saarenpää, Ahti/Eder, Stefan/Zanol, Jakov/Schmautzer, Felix/Kummer, Franz/Hanke, Philip (Hrsg.), Recht DIGITAL – 25 Jahre IRIS, Tagungsband des 25. Internationalen Rechtsinformatik Symposions IRIS 2022, Bern 2022; darin: *Mielke, Bettina/Wolff, Christian*, Der strukturierte Parteivortrag im Zivilprozess, S. 195-204 (zit. als *Mielke/Wolff*, Der strukturierte Parteivortrag im Zivilprozess, in: Schweighofer/Saarenpää/Eder/Zanol/Schmautzer/Kummer/Hanke (Hrsg.), Recht DIGITAL – 25 Jahre IRIS, 2022, S. *Seitenzahl*).

VI. Aufsätze

AlHidaifi, Saleh Mohamed/Asghar, Muhammad Rizwan/Ansari, Imran Shafique, A Survey on Cyber Resilience: Key Strategies, Research Challenges, and Future Directions, ACM Computing Surveys 56 (2024), Nr. 8, Art. 196.

Banks, Isabella, Judges-in-the-loop? Judicial involvement in human oversight of high-risk decision support systems under the EU AI Act, International Journal of Law and Information Technology 34 (2026), eaag001.

Beretas, Christos P., Information Systems Security, Detection and Recovery from Cyber Attacks, Universal Library of Engineering Technology 1 (2024), Nr. 1, 27.

Berlit, Uwe, Richterliche Unabhängigkeit und elektronische Akte, JurPC Web-Dok. 77/2012.

Bomhard, David/Etzkorn, Philipp, Impulse zum Entwurf der KI-Verordnung, ZdiW 2022, 159.

Bovens, Mark, Analysing and Assessing Accountability: A Conceptual Framework, European Law Journal 13 (2007), 447.

Brodowski, Dominik/Grabmair, Matthias/Langer, Markus/Paschke, Anne/Rachutz, Sarah/Reuß, Philipp M./Treffers, Theresa/Welpe, Isabell M./Wörner, Liane, Überlegungen zu Gelingensbedingungen für den akzeptablen Einsatz von Künstlicher Intelligenz im Justizsystem, RDi 2026, 585.

Buçinca, Zana/Malaya, Maja Barbara/Gajos, Krzysztof Z., To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making, Proceedings of the ACM on Human-Computer Interaction 5 (2021), CSCW1, Art. 188.

Bystranowski, Piotr/Janik, Bartosz/Próchnicki, Maciej/Skórska, Paulina, Anchoring Effect in Legal Decision-Making: A Meta-Analysis, Law and Human Behavior 45 (2021), 1.

Carnat, Irina, Addressing the risks of generative AI for the judiciary: The accountability framework(s) under the EU AI Act, Computer Law & Security Review 55 (2024), Art. 106067.

Caroei, Asya S./Brodowski, Dominik/Grabmair, Matthias/Koçak, Aleyna/Kolbe, Moritz/Olheide, Jens/Paschke, Anne/Rachut, Sarah/Reuß, Philipp M./Sigmund, Johannes/Treffers, Theresa/Uth, Richard/Welpe, Isabell M./Wörner, Liane/Zettelmeier, Anne/Langer, Markus, KI-Einsatz im deutschen Justizsystem – Ergebnisse einer empirischen Untersuchung, MMR 2026, 619.

Choi, Jonathan H., Off-The-Shelf Large Language Models Are Unreliable Judges, Journal of Empirical Legal Studies 2026, 1.

Cox, Louis Anthony, Improving Causal Inferences in Risk Analysis, Risk Analysis 33 (2013), 1762.

Cox, Louis Anthony/Ricci, Paolo F., Causation in risk assessment and management: models, inference, biases, and a microbial risk-benefit case study, Environment International 31 (2005), 377.

Danziger, Shai/Levav, Jonathan/Avnaim-Pesso, Liora, Extraneous factors in judicial decisions, Proceedings of the National Academy of Sciences 108 (2011), 6889.

Dharanikota, Harini/Howie, Emma/Hope, Lorraine/Wigmore, Stephen J./Skipworth, Richard J. E./Yule, Steven, Debiasing Judgements Using a Distributed Cognition Approach: A Scoping Review of Technological Strategies, Human Factors 67 (2025), 525.

Dong, Mengchen/Bonnefon, Jean-François/Rahwan, Iyad, Heterogeneous preferences and asymmetric insights for AI use among welfare claimants and non-claimants, Nature Communications 16 (2025), Art. 6973.

Dötterl, Sebastian, „Unterstützen ja, Ersetzen nein”? Menschliche Entscheidungshoheit beim KI-Einsatz in der Justiz, RDi 2026, 414.

Edmond, Gary/Martire, Kristy A., Just Cognition: Scientific Research on Bias and Some Implications for Legal Procedure and Decision-Making, The Modern Law Review 82 (2019), 633.

Elliott, Matthew/Golub, Benjamin, Networks and Economic Fragility, Annual Review of Economics 14 (2022), 665.

Fink, Melanie, Human Oversight under Article 14 of the EU AI Act, 2025, abrufbar unter https://doi.org/10.2139/ssrn.5147196 (Abruf am 25.10.2026, 14:02 Uhr).

Gaube, Susanne/Langer, Markus/Miller, Tim/Baum, Kevin/Dachselt, Raimund/Feit, Anna Maria/Gadiraju, Ujwal/Kaur, Harmanpreet/Keane, Mark T./Landers, Richard/Laux, Johann/Liao, Q. Vera/Lim, Brian/Onnasch, Linda/Schrills, Tim/Sonenberg, Liz/Tan, Chenhao/Tintarev, Nava/Xiao, Ziang/Zhang, Hanwei, Keeping an Eye on AI: A Framework for Effective Human Oversight of AI Systems, 2026, abrufbar unter https://doi.org/10.48550/arXiv.2605.16278 (Abruf am 24.10.2026, 23:00 Uhr).

Gebru, Timnit/Morgenstern, Jamie/Vecchione, Briana/Vaughan, Jennifer Wortman/Wallach, Hanna/Daumé III, Hal/Crawford, Kate, Datasheets for Datasets, abrufbar unter https://doi.org/10.48550/arXiv.1803.09010 (Abruf am 23.09.2026, 20:42 Uhr).

Goddard, Kate/Roudsari, Abdul/Wyatt, Jeremy C., Automation bias: a systematic review of frequency, effect mediators, and mitigators, Journal of the American Medical Informatics Association 19 (2012), 121.

Green, Ben, The flaws of policies requiring human oversight of government algorithms, Computer Law & Security Review 45 (2022), Art. 105681.

Green, Ben/Chen, Yiling, Disparate Interactions: An Algorithm-in-the-Loop Analysis of Fairness in Risk Assessments, Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19), 2019, 90.

Griesel, Moritz/Brorsen, Hans, Kennzeichnungspflichten für synthetische Medieninhalte – Rechtliche Anforderungen und technische Realität, MMR 2026, 477.

Gsell, Beate/Pfündl, Bodo, KI im Zivilprozess und menschliche Entscheidungsverantwortung – Sicherung richterlicher Unabhängigkeit durch Parteiherrschaft und kontradiktorisches Verfahren, RDi 2026, 457.

Guijarro Santos, Victoria, Nicht besser als nichts – Ein Kommentar zum KI-Verordnungsentwurf, ZfDR 2023, 23.

Guthrie, Chris/Rachlinski, Jeffrey J./Wistrich, Andrew J., Judging by Heuristic: Cognitive Illusions in Judicial Decision Making, Judicature 86 (2002), 44.

Hacker, Philipp/Cordes, Johann/Berz, Amelie, Transparenz generativer KI – Rechtliche Rahmenbedingungen und technische Möglichkeiten, GRUR 2024, 1777.

Haldane, Andrew G./May, Robert M., Systemic risk in banking ecosystems, Nature 469 (2011), 351.

Hofer, Matthias/Kirchmair, Verena, KI-Kompetenz in der Praxis: Compliance-Strategien für Unternehmen im Lichte der KI-VO, ZfPC 2025, 270.

Hoffmann, Sina Marie, Zugangshürden durch die Digitalisierung des Zivilprozesses?, RDi 2022, 76.

Horn, Jakob, Die vollautomatische zivilgerichtliche Entscheidung, ZZP 136 (2023), 459.

Krauss, Daniel A./Cook, Gabriel I./Klapatch, Lukas, Risk assessment communication difficulties: An empirical examination of the effects of categorical versus probabilistic risk communication in sexually violent predator decisions, Behavioral Sciences & the Law 36 (2018), 532.

Kuhlmann, Nico, Die Nutzung von KI durch Juristen – ein Überblick, NJW 2026, 2774.

Magesh, Varun/Surani, Faiz/Dahl, Matthew/Suzgun, Mirac/Manning, Christopher D./Ho, Daniel E., Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools, Journal of Empirical Legal Studies 22 (2025), 216.

Manakul, Potsawee/Liusie, Adian/Gales, Mark, SelfCheckGPT: Zero-Resource Black-Box Hallucination Detection for Generative Large Language Models, roceedings of the 2023 Conference on Empirical Methods in Natural Language Processing 2023, 9004.

Martini, Mario/Nink, David, Wenn Maschinen entscheiden … – vollautomatisierte Verwaltungsverfahren und der Persönlichkeitsschutz, NVwZ-Extra 10/2017, 1.

Matthias, Andreas, The responsibility gap: Ascribing responsibility for the actions of learning automata, Ethics and Information Technology 6 (2004), 175.

Meinel, Philip/Baum, Kevin/Lauber-Rönsberg, Anne/Stenzel, Laura/Langer, Markus, Human Oversight of AI Agents under the EU AI Act, 2025, abrufbar unter https://dx.doi.org/10.2139/ssrn.7500959 (Abruf am 21.10.2026, 22:52 Uhr).

Meller-Hannich, Caroline, Beschleunigtes Online-Verfahren und Verfahrensgrundsätze – Eine sorgfältige Neuinterpretation könnte Verfahrensrechte stärken, AnwBl 2021, 288.

Meyer, Stephan, Künstliche Intelligenz und die Rolle des Rechts für Innovation – Rechtliche Rationalitätsanforderungen an zukünftige Regulierung, ZRP 2018, 233.

Mielke, Bettina/Wolff, Christian, Neue Wege im Zivilprozess: Das digitale Basisdokument – Einblick in das bundesweit erste Reallabor im Rechtswesen, RDi 2024, 361.

Miltgen, Caroline Lancelot/Smith, H. Jeff, Exploring information privacy regulation, risks, trust, and behavior, Information & Management 52 (2015), Nr. 6, 741.

Min, Sewon/Krishna, Kalpesh/Lyu, Xinxi/Lewis, Mike/Yih, Wen-tau/Koh, Pang Wei/Iyyer, Mohit/Zettlemoyer, Luke/Hajishirzi, Hannaneh, FActScore: Fine-grained Atomic Evaluation of Factual Precision in Long Form Text Generation, Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing 2023, 12076.

Molavi Vasse'i, Ramak, Watermarking von KI-generierten Inhalten als regulatorisches Instrument – Risiken synthetischer Inhalte und maschinenlesbare Kennzeichnungsmethoden –, RDi 2024, 406.

Mulligan, Deirdre K./Bamberger, Kenneth A., Procurement as Policy: Administrative Process for Machine Learning, Berkeley Technology Law Journal 34 (2019), 773.

Nagl, Sebastian/Mayrhofer, Ann-Kristin/Heidebach, Martin/Koçak, Aleyna/Zettelmeier, Anne/Breu, Elly/Greiner, Angelina/Milijas, Sofija/Grabmair, Matthias, BenGER: Benchmarking LLM Systems on Subsumption-Based Legal Reasoning in German Law, 2026, abrufbar unter https://arxiv.org/abs/2605.28183 (Abruf am 25.10.2026, 14:05 Uhr).

Notz, Anna von, Die vernetzte Justiz, JZ 2017, 607.

Omar, Mahmud/Sorin, Vera/Collins, Jeremy D./Reich, David/Freeman, Robert/Gavin, Nicholas/Alexander, Charney/Stump, Lisa/Bragazzi, Nicol Luigi/Nadkarni, Girish N./Klang, Eyal, Multi-model assurance analysis showing large language models are highly vulnerable to adversarial hallucination attacks during clinical decision support, Communications Medicine, 5(1), 2025, Artikel 330.

Oster, Jan, Kennzeichnungspflichten für KI-generierte Inhalte nach Art. 50 KI-VO, ZfDR 2026, 163.

Piccardi, Carlo/Tajoli, Lucia, Complexity, centralization, and fragility in economic networks, PLOS ONE 13 (2018), e0208265.

Pillai, Vinayak, Enhancing transparency and understanding in AI decision-making processes, Iconic Research and Engineering Journals 8 (2024), Nr. 1, 168.

Radke, Holger, Datenhaltung und Datenadministration der Justiz und richterliche Unabhängigkeit, jM 2016, 8.

Reuß, Philipp M., Kann KI helfen, höchstrichterliche Rechtsprechung für Bürgerinnen und Bürger zugänglicher zu machen?, PdR 2026, 77.

Riehm, Thomas/Valerius, Brian, KI-Halluzinationen im Zivilprozess, RDi 2025, 612.

Rühl, Giesela, Digitale Justiz, oder: Zivilverfahren für das 21. Jahrhundert, JZ 2020, 809.

Santoni de Sio, Filippo/Mecacci, Giulio, Four Responsibility Gaps with Artificial Intelligence: Why they Matter and How to Address them, Philosophy & Technology 34 (2021), 1057.

Santoni de Sio, Filippo/van den Hoven, Jeroen, Meaningful Human Control over Autonomous Systems: A Philosophical Account, Frontiers in Robotics and AI 5 (2018), Art. 15.

Schkade, David A./Kahneman, Daniel, Does Living in California Make People Happy? A Focusing Illusion in Judgments of Life Satisfaction, Psychological Science 9 (1998), 340.

Schneider, Sven H./Wundenberg, Malte/Sharaf, Samy G., Künstliche Intelligenz und interne Untersuchungen, ZHR 189 (2025), 259.

Spindler, Gerald, Der Vorschlag der EU-Kommission für eine Verordnung zur Regulierung der Künstlichen Intelligenz (KI-VO-E), CR 2021, 361.

Stafford, Tom, Biases in decision making, Tribunals, Winter 2017, 19.

Sterz, Sarah/Baum, Kevin/Biewer, Sebastian/Hermanns, Holger/Lauber-Rönsberg, Anne/Meinel, Philip/Langer, Markus, On the Quest for Effectiveness in Human Oversight: Interdisciplinary Perspectives, 2024, abrufbar unter https://doi.org/10.48550/arXiv.2404.04059 (Abruf am 22.10.2026, 20:38 Uhr).

Stolfo, Alessandro, Groundedness in Retrieval-augmented Long-form Generation: An Empirical Study, Findings of the Association for Computational Linguistics: NAACL 2024, 1537.

Stürner, Michael, Künstliche Intelligenz im Zivilprozess – welche Perspektiven bietet die E-Justiz-Strategie 2024–2028?, GPR 2025, 62.

Sundararajan, Barkavi/Sripada, Somayajulu/Reiter, Ehud, Input Matters: Evaluating Input Structure’s Impact on LLM Summaries of Sports Play-by-Play, Proceedings of the 18th International Natural Language Generation Conference 2025, 795.

Suresh, Harini/Guttag, John, A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle, Proceedings of the 1st ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '21), Association for Computing Machinery 2021, Article 17.

Tonzer, Lena, Cross-border interbank networks, banking risk and contagion, Journal of Financial Stability 18 (2015), 19.

Vogelgesang, Stephanie, Datensicherheit und IT-Sicherheit in der Justiz, jM 2018, 2.

Wachter, Martin/Leeb, Christina-Maria, KI-Systeme in der Rechtspflege, RDi 2024, 440.

Wang, Ruotong/Harper, F. Maxwell/Zhu, Haiyi, Factors Influencing Perceived Fairness in Algorithmic Decision-Making: Algorithm Outcomes, Development Procedures, and Individual Differences, Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, 1.

Weber, Franziska, Gegenwärtige Verbraucherrechtsfälle und Bedarf an staatlicher Rechtsdurchsetzung, VuR 2013, 323.

Weinshall-Margel, Keren/Shapard, John, Overlooked factors in the analysis of parole decisions, Proceedings of the National Academy of Sciences 108 (2011), E833.

Wischmeyer, Thomas, Regulierung intelligenter Systeme, AöR 143 (2018), 1.

Wistrich, Andrew J./Guthrie, Chris/Rachlinski, Jeffrey J., Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding, University of Pennsylvania Law Review 153 (2005), 1251.

VII. Internetdokumente

Amos, Maximilian, NRW will Arbeitsgerichte zusammenstutzen, abrufbar unter https://rsw.beck.de/aktuell/daily/meldung/detail/diskussionspapier-nrw-arbeitsgerichte-strukturreform-zusammenlegung (Abruf am 18.03.2026, 14:59 Uhr).

Amtsgericht Euskirchen, Grundsätzliches und Zuständigkeit, abrufbar unter https://www.ag-euskirchen.nrw.de/behoerde/zema/grundsaetzliches-zustaendigkeit/index.php (Abruf am 26.02.2026, 15:53 Uhr).

Anthropic, Disrupting the first reported AI-orchestrated cyber espionage campaign, 13.11.2025, abrufbar unter https://www.anthropic.com/news/disrupting-AI-espionage (Abruf am 25.10.2026, 14:09 Uhr).

Anthropic, Wie Claude KI-generierte Inhalte kennzeichnet, abrufbar unter https://support.claude.com/de/articles/16266773-wie-claude-ki-generierte-inhalte-kennzeichnet (Abruf am 01.09.2026, 14:59 Uhr).

Bayerisches Landesamt für Datenschutzaufsicht, KI & Datenschutz, abrufbar unter https://www.lda.bayern.de/de/ki.html (Abruf am 14.04.2026, 17:05 Uhr).

Bölling, Noëlle, Anthropic erklärt, wie das neue Wasserzeichen in Claude genau funktioniert, abrufbar unter https://t3n.de/news/anthropic-erklaert-wie-das-neue-wasserzeichen-in-claude-genau-funktioniert-1758332/ (Abruf am 01.09.2026, 15:03 Uhr).

Bundesgerichtshof, Anfahrt zum Bundesgerichtshof in Karlsruhe, abrufbar unter https://www.bundesgerichtshof.de/DE/Service/Kontakt/AnfahrtKarlsruhe/anfahrtKarlsruhe_node.html (Abruf am 19.10.2026, 16:59 Uhr).

Bundesgerichtshof, Anfahrt zur Außenstelle des Bundesgerichtshofs in Karlsruhe, abrufbar unter https://www.bundesgerichtshof.de/DE/Service/Kontakt/AnfahrtKarlsruheRQ/anfahrtKarlsruheRQ_node.html (Abruf am 19.10.2026, 16:59 Uhr).

Bundesgerichtshof, Anfahrt zur Dienststelle Leipzig des Bundesgerichtshofs, abrufbar unter

https://www.bundesgerichtshof.de/DE/Service/Kontakt/AnfahrtLeipzig/anfahrtLeipzig_node.html (Abruf am 19.10.2026, 16:59 Uhr).

Bundesministerium der Justiz und für Verbraucherschutz, Gerichtsentscheidungen durch KI leichter veröffentlichen, abrufbar unter https://www.bmjv.de/DE/themen/digitales/digitalisierung_justiz/digitalisierungsinitiative/laendervorhaben/_doc/artikel_vorhaben_14_ALeKS.html (Abruf am 31.03.2026, 19:45 Uhr).

Bundesministerium der Justiz und für Verbraucherschutz, Massenverfahren mit KI schneller und effizienter bearbeiten, abrufbar unter https://www.bmjv.de/DE/themen/digitales/digitalisierung_justiz/digitalisierungsinitiative/laendervorhaben/_doc/artikel_vorhaben_13_MAKI.html (Abruf am 22.09.2026, 18:07 Uhr).

Bundesministerium für Digitales und Staatsmodernisierung, Digitale Barrierefreiheit, Stand: 17.06.2026, abrufbar unter https://www.barrierefreiheit-dienstekonsolidierung.bund.de/Webs/PB/DE/barrierefreie_it/digitale-barrierefreiheit/digitale-barrierefreiheit-node.html (Abruf am 23.09.2026, 20:16 Uhr).

Bundesministerium für Digitales und Staatsmodernisierung, Lexikon, Stichwort „Medienbruch“, abrufbar unter https://www.digitale-verwaltung.de/Webs/DV/DE/aktuelles-service/faq-lexikon/glossar/functions/ozg-lexikon.html?lv3=20348760&lv2=20339810#doc20348760 (Abruf am 04.06.2026, 19:52 Uhr).

Der Bundesbeauftragte für den Datenschutz und die Informationsfreiheit, Das Standard-Datenschutzmodell, abrufbar unter https://www.bfdi.bund.de/DE/Fachthemen/Inhalte/Technik/SDM.html (Abruf am 08.04.2026, 17:28 Uhr).

Der Landesbeauftragte für Datenschutz und Informationsfreiheit Mecklenburg-Vorpommern, Das Standard-Datenschutzmodell, abrufbar unter https://www.datenschutz-mv.de/datenschutz/datenschutzmodell/ (Abruf am 08.04.2026, 17:30 Uhr).

Der Landesbeauftragte für den Datenschutz Niedersachsen, Muss-Listen zur Datenschutz-Folgenabschätzung, abrufbar unter https://www.lfd.niedersachsen.de/dsgvo/liste_von_verarbeitungsvorgangen_nach_art_35_abs_4_ds_gvo/muss-listen-zur-datenschutz-folgenabschatzung-179663.html (Abruf am 09.04.2026, 22:25 Uhr).

DeutscheAnwaltAkademie Gesellschaft für Aus- und Fortbildung sowie Serviceleistungen mbH, Fachanwaltsordnung, abrufbar unter https://www.anwaltakademie.de/fachanwaltslehrgaenge/fachanwaltsordnung (Abruf am 04.05.2026).

Deutscher Bundestag, Gesetz zur Einführung der Zulassungsrevision gegen Berufungsurteile der Landgerichte in Zivilsachen, abrufbar unter https://dip.bundestag.de/vorgang/gesetz-zur-einf%C3%BChrung-der-zulassungsrevision-gegen-berufungsurteile-der-landgerichte-in/305207 (Abruf am 25.02.2026, 17:47 Uhr).

Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, DFKI, Re-Engineering, abrufbar unter https://werner.dfki.de/technologydetail/2e453901-e173-477e-bbe3-63328323e349 (Abruf am 09.09.2026, 20:57 Uhr).

Die Zeit, Ein Bagger, ein Kabel – und die Justiz steht still, abrufbar unter https://www.zeit.de/news/2026-06/10/kabel-beschaedigt-landesweite-it-stoerung-trifft-gerichte (Abruf am 13.07.2026).

dpa/sh/LTO-Redaktion, Ost-Länder wollen Gerichte zusammenlegen, abrufbar unter https://www.lto.de/recht/nachrichten/n/justiz-ost-laender-wollen-gerichte-zusammenlegen (Abruf am 18.03.2026, 12:36 Uhr).

dpa/xp/LTO-Redaktion, „Wir stehen mit dem Rücken zur Wand“, abrufbar unter https://www.lto.de/recht/justiz/j/nordrhein-westfalen-sozialgerichte-wegen-ki-schriftsaetzen-ueberlastet (Abruf am 18.08.2026, 17:45 Uhr).

Duden, Bias, abrufbar unter https://www.duden.de/rechtschreibung/Bias (Abruf am 30.07.2026, 14:08 Uhr).

Eurostat, Niveau der digitalen Kompetenzen von Einzelpersonen (2015-2019), abrufbar unter https://ec.europa.eu/eurostat/databrowser/view/isoc_sk_dskl_i$DV_317/default/table?lang=de.

Fuhrmann, Marvin, Claude-Wasserzeichen entfernen: Dieses Tool verspricht Abhilfe – aber funktioniert es wirklich?, abrufbar unter https://t3n.de/news/claude-wasserzeichen-entfernen-tool-anthropic-synthid-1758888/ (Abruf am 01.09.2026, 15:11 Uhr).

Georg-August-Universität Göttingen, Neuer Zertifikatsstudiengang: KI und Legal Tech, abrufbar unter https://uni-goettingen.de/de/neuer+zertifikatsstudiengang%3A+ki+und+legal+tech/698161.html (Abruf am 24.10.2026, 17:04 Uhr).

Georg-August-Universität Göttingen, Umgang mit KI-Modellen in Studium und Lehre, abrufbar unter https://uni-goettingen.de/de/674738.html (Abruf am 20.04.2026, 19:52 Uhr).

Google Threat Intelligence Group, GTIG AI Threat Tracker: Advances in threat actor usage of AI tools, November 2025, abrufbar unter https://cloud.google.com/blog/topics/threat-intelligence/threat-actor-usage-of-ai-tools (Abruf am 25.10.2026, 14:11 Uhr).

Irskens, Gesine, KI in der niedersächsischen Justiz, Anwaltspraxis Magazin vom 09.02.2025, abrufbar unter https://anwaltspraxis-magazin.de/fachbeitraege/elektronischer-rechtsverkehr/2025/02/09/b-ki-in-der-niedersaechsischen-justiz-80/ (Abruf am 04.11.2026, 19:48 Uhr).

ITD, Organisation, abrufbar unter https://www.justiz.nrw.de/Gerichte_Behoerden/zentraler_dienstleister/ah20_organisation (Abruf am 28.01.2026).

Knobloch, Andreas, Claude: KI-Chatbot für Cyberangriff auf mexikanische Regierung genutzt, abrufbar unter https://www.heise.de/news/Claude-KI-Chatbot-fuer-Cyberangriff-auf-mexikanische-Regierung-genutzt-11190396.html (Abruf am 23.03.2026, 21:45 Uhr).

Kosinski, Matthew, So verhindern Sie Prompt-Injection-Angriffe, abrufbar unter https://www.ibm.com/de-de/think/insights/prevent-prompt-injection (Abruf am 23.04.2026, 21:34 Uhr).

Kuhlmann, Nico, Prompt Injections – Betrügen für Fortgeschrittene, abrufbar unter https://www.lto.de/recht/hintergruende/h/prompt-injections-erklaerung-ist-das-strafbar (Abruf am 23.04.2026, 21:16 Uhr).

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, Deskilling, abrufbar unter https://www.lernen-digital.nrw/glossary-term/100 (Abruf am 24.09.2026, 15:42 Uhr).

Niedersächsisches Justizministerium, Künstliche Intelligenz und Automation in der Justiz, abrufbar unter https://www.mj.niedersachsen.de/startseite/kunstliche_intelligenz_und_automation_in_der_justiz/kunstliche-intelligenz-und-automation-in-der-justiz-246705.html (Abruf am 31.03.2026, 19:41 Uhr).

Norton Rose Fulbright, AI in litigation: Update on Gen AI sanctions in 2026, abrufbar unter https://www.nortonrosefulbright.com/en-us/knowledge/publications/792d8bf3/ai-in-litigation-update-on-gen-ai-sanctions-in-2026 (Abruf am 25.10.2026, 14:14 Uhr).

OpenAI, Disrupting malicious uses of AI: October 2025, 07.10.2025, abrufbar unter https://openai.com/global-affairs/disrupting-malicious-uses-of-ai-october-2025 (Abruf am 25.10.2026, 14:18 Uhr).

OpenAI, Prompt Injections verstehen: eine neuartige Sicherheitsherausforderung, abrufbar unter https://openai.com/de-DE/index/prompt-injections/ (Abruf am 23.04.2026, 21:19 Uhr).

Reuß, Philipp M., Göttinger Forschungsstelle für e-Justiz, Automation und Digitalisierung (GOeJUSTICE), abrufbar unter https://reusz.eu/goejustice (Abruf am 05.08.2026, 21:30 Uhr).

Sieb, Antje, OpenAI veröffentlicht weitere KI-Zwischenfälle, abrufbar unter https://www.tagesschau.de/ausland/amerika/openai-kuenstliche-intelligenz-100.html (Abruf am 26.09.2026, 22:00 Uhr).

Society for Risk Analysis, Society for Risk Analysis Glossary, 2018, abrufbar unter https://www.sra.org/risk-analysis-introduction/risk-analysis-glossary/ (Abruf am 25.10.2026, 14:20 Uhr).

Stabsstelle LBIT – Landeskompetenzzentrum für barrierefreie Informationstechnik und Durchsetzungs- und Überwachungsstelle, Zielgruppe Hören, abrufbar unter https://lbit.hessen.de/landesbeauftragte-fuer-barrierefreie-it/zielgruppen-barrierefreier-gestaltung/zielgruppe-hoeren (Abruf am 23.09.2026, 20:35 Uhr).

Stabsstelle LBIT – Landeskompetenzzentrum für barrierefreie Informationstechnik und Durchsetzungs- und Überwachungsstelle, Zielgruppe Sehen, abrufbar unter https://lbit.hessen.de/landesbeauftragte-fuer-barrierefreie-it/zielgruppen-barrierefreier-gestaltung/zielgruppe-sehen (Abruf am 23.09.2026, 20:33 Uhr).

Stabsstelle LBIT – Landeskompetenzzentrum für barrierefreie Informationstechnik und Durchsetzungs- und Überwachungsstelle, Zielgruppe Verstehen & Interagieren, abrufbar unter https://lbit.hessen.de/landesbeauftragte-fuer-barrierefreie-it/zielgruppen-barrierefreier-gestaltung/zielgruppe-interaktion-verstaendnis (Abruf am 23.09.2026, 20:36 Uhr).

TITAN-Forschungsverbund (Technologische Intelligenz zur Transformation, Automatisierung und Nutzerorientierung des Justizsystems) (Prof. Grabmair/TU Braunschweig), Darstellung Akzeptabilität, abrufbar unter https://www.tu-braunschweig.de/recht/digitl-hub/projekte/titan/vision/zielvorstellung (Abruf am 11.08.2026, 21:34 Uhr).

ZDF, „Claude Mythos“ von Anthropic: Neues KI-Modell: BSI zeigt sich besorgt, abrufbar unter https://www.zdfheute.de/politik/deutschland/ki-anthropic-claude-mythos-schwachstellen-software-bsi-100.html (Abruf am 23.04.2026, 21:53 Uhr).

Zenthöfer, Jochen, Der peinliche Copy-Paste-Fehler hessischer Richter mit ChatGPT, abrufbar unter https://www.faz.net/aktuell/feuilleton/medien-und-film/medienpolitik/der-peinliche-copy-paste-fehler-hessischer-richter-mit-chatgpt-201105297.html (Abruf am 18.08.2026, 18:01 Uhr).

Zentraler IT-Betrieb Niedersächsische Justiz, Impressum, abrufbar unter https://zib.niedersachsen.de/startseite/service_neu/impressum-197077.html (Abruf am 06.07.2026).

VIII. Sonstiges

Arbeitsgruppe „Einsatz von KI und algorithmischen Systemen in der Justiz“, Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier zur 74. Jahrestagung der Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs vom 23. bis 25. Mai 2022 in Rostock, abrufbar unter https://www.justiz.bayern.de/media/images/behoerden-und-gerichte/oberlandesgerichte/nuernberg/einsatz_von_ki_und_algorithmischen_systemen_in_der_justiz.pdf (Abruf am 15.04.2026, 18:58 Uhr).

Artikel-29-Datenschutzgruppe, WP 248 rev.01, abrufbar unter https://www.datenschutzkonferenz-online.de/media/wp/20171004_wp248_rev01.pdf (Abruf am 14.04.2026, 20:28 Uhr).

Artikel-29-Datenschutzgruppe, WP 251 rev.01, Guidelines on automated individual decision-making and profiling for the purposes of Regulation 2016/679, abrufbar unter https://www.edpb.europa.eu/documents/guideline/automated-decision-making-and-profiling_en (Abruf am 25.10.2026, 14:23 Uhr).

Artikel-29-Datenschutzgruppe, WP 260 rev.01, abrufbar unter https://www.lfd.niedersachsen.de/startseite/infothek/edsa_leitlinien/vom-edsa-bestaetigte-leitlinien-158964.html (Abruf am 16.09.2026, 23:30 Uhr).

Bitkom e. V., Umsetzungsleitfaden zur KI-Verordnung, Version 2.0, abrufbar unter https://www.bitkom.org/Bitkom/Publikationen/Umsetzungsleitfaden-zur-KI-Verordnung (Abruf am 18.08.2026, 20:42 Uhr).

Bontcheva, Kalina/Bechmann, Anja et al., Verhaltenskodex für die Transparenz von KI-generierten Inhalten, „Code of Practice on Transparency of AI-Generated Content“, abrufbar unter https://digital-strategy.ec.europa.eu/de/policies/code-practice-ai-generated-content (Abruf am 03.09.2026, 15:58 Uhr).

Borchert, Sarah Jane/Bredowski, Katja/Delor, André/Eppig, Petra/Göhner, Kerstin/Jenke, Marion/Jessen-Narr, Frauke/Krimitzas, Julia/Lange-Kniep, Annika/Nägele, Katrin/Nietzio, Annika/Nolte, Henrik/Rhensius, Ruben/Schnepel, Ole, Kodex für den Einsatz von Künstlicher Intelligenz für Leichte-Sprache-Texte, 2025, abrufbar unter https://www.dicv-limburg.de/derverband/wissenswertes/leichte-sprache/kodex-fuer-ki-gestuetzte-uebersetzungen/kodex-fuer-ki-gestuetzte-uebersetzungen (Abruf am 13.10.2026, 20:16 Uhr).

Bundesamt für Sicherheit in der Informationstechnik (BSI), A5 – Horizontales Trustworthiness-Basismodul, Community-Draft, Version 0.9, Stand 30.06.2026, abrufbar unter

https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KI/A5/A5_Horizontales-Trustworthiness-Basismodul.pdf (Abruf am 25.10.2026, 14:27 Uhr).

Bundesministerium für Arbeit und Soziales/Bundesfachstelle Barrierefreiheit, Leitfaden „Umsetzung digitaler Barrierefreiheit“, Februar 2026, abrufbar unter https://www.esf-bw.de/esf-plus-in-bw/querschnittsziele-und-grundlegende-voraussetzungen/chancengleichheit-und-nichtdiskriminierung (Abruf am 23.09.2026, 21:22 Uhr).

Bundesministerium für Arbeit und Soziales, Leitlinien für die Anwendung des Barrierefreiheitsstärkungsgesetzes, abrufbar unter https://www.bundesfachstelle-barrierefreiheit.de/SharedDocs/Downloads/DE/Externe-Veroeffentlichungen/bmas-leitlinien-bfsg (Abruf am 23.09.2026, 20:12 Uhr).

Deutscher Bundestag, Fachbereich Europa, Zu rechtlichen Vorgaben der Verordnung über Künstliche Intelligenz für den Einsatz von KI in Behörden der EU-Mitgliedstaaten, Infobrief EU 6 - 3010 - 032/25, 07.07.2025.

E-Justice-Rat, Strategie für den Einsatz von Künstlicher Intelligenz in der Justiz, Version 1.0, April 2025, abrufbar unter https://justiz.de/laender-bund-europa/e_justice_rat/beschluesse/index.php (Abruf am 23.09.2026, 21:17 Uhr).

Europäische Kommission, Entwurf der Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme gemäß Artikel 50 der KI-VO, abrufbar unter https://digital-strategy.ec.europa.eu/en/library/draft-guidelines-implementation-transparency-obligations-certain-ai-systems-under-article-50-ai-act (Abruf am 13.05.2026, 16:15 Uhr).

Europäische Kommission, Leitlinien zur Umsetzung der Transparenzpflichten für bestimmte KI-Systeme nach Artikel 50 der KI-Verordnung „Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the ‘AI Act’)“, abrufbar unter https://digital-strategy.ec.europa.eu/de/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems (Abruf am 31.08.2026, 20:51 Uhr).

Europäische Kommission, Pressemitteilung 20.07.2026, Kommission veröffentlicht Leitlinien zu Transparenzpflichten für Anbieter und Betreiber bestimmter KI-Systeme, abrufbar unter https://ec.europa.eu/commission/presscorner/detail/de/ip_26_1653 (Abruf am 31.08.2026, 20:51 Uhr).

European Commission for the Efficiency of Justice (CEPEJ), European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment, 2018, abrufbar unter https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c (Abruf am 24.04.2026, 14:58 Uhr).

European Commission for the Efficiency of Justice (CEPEJ), Guidelines on the use of generative AI for courts, CEPEJ(2025)18Final, 2025, abrufbar unter https://rm.coe.int/cepej-2025-18final-en-draft-guidelines-on-the-use-of-generative-ai-for/48802a4ad1 (Abruf am 25.10.2026, 14:29 Uhr).

Greger, Reinhard, Stellungnahme zum Entwurf eines Gesetzes zur Änderung der Gesetzes, betreffend die Einführung der Zivilprozessordnung (BT-Drucksache 19/1686) vom 08.05.2018, abrufbar unter https://www.bundestag.de/resource/blob/554494/c352b8830ab4c3abe77e3c3d2b32aa94/greger-data.pdf (Abruf am 17.03.2026, 13:58 Uhr).

International Organization for Standardization (ISO), ISO 31000:2018 – Risk management – Guidelines, Genf 2018.

International Organization for Standardization/International Electrotechnical Commission (ISO/IEC), ISO/IEC 42001:2023 – Information technology – Artificial intelligence – Management system, Genf 2023.

Ministerium der Justiz des Landes Nordrhein-Westfalen, Arbeitsgerichtsbarkeit der Zukunft in Nordrhein-Westfalen, Diskussionspapier für den weiteren Beteiligungsprozess, abrufbar unter https://www.justiz.nrw.de/sites/default/files/2025-11/2025-11-11%20Diskussionspapier.pdf (Abruf am 18.03.2026, 15:28 Uhr).

National Cyber Security Centre u. a., Guidelines for secure AI system development, abrufbar unter https://www.ncsc.gov.uk/sites/default/files/documents/Guidelines-for-secure-AI-system-development.pdf (Abruf am 23.04.2026, 13:14 Uhr).

National Institute of Standards and Technology (NIST), Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST AI 100-1, Gaithersburg 2023, abrufbar unter https://doi.org/10.6028/NIST.AI.100-1 (Abruf am 25.10.2026, 14:32 Uhr).

Partnership on AI, Building a Glossary for Synthetic Media Transparency Methods, Part 1: Indirect Disclosure, 2023.

Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs, Beschluss vom Mai 2026 zum Grundlagenpapier 2026 zum Einsatz von KI und algorithmischen Systemen in der Justiz, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 20:49 Uhr).

Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs, Einsatz von KI und algorithmischen Systemen in der Justiz – Grundlagenpapier 2026, abrufbar unter https://oberlandesgericht-celle.niedersachsen.de/startseite/rechtsprechung/zivilprozess_der_zukunft_tagungsband/justiz_der_zukunft/aktualisiertes-grundlagenpapier-zum-einsatz-von-ki-und-algorithmischen-systemen-in-der-justiz-250922.html (Abruf am 03.06.2026, 14:17 Uhr).

Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs (Auftraggeber), Thesenpapier der Arbeitsgruppe „Modernisierung des Zivilprozesses“, abrufbar unter https://www.justiz.bayern.de/media/images/behoerden-und-gerichte/oberlandesgerichte/nuernberg/thesenpapier_der_arbeitsgruppe.pdf (Abruf am 26.02.2026, 13:01 Uhr).

Präsidentinnen und Präsidenten der Oberlandesgerichte, des Kammergerichts, des Bayerischen Obersten Landesgerichts und des Bundesgerichtshofs, Münchener Thesen zum Zivilprozess der Zukunft, abrufbar unter https://www.olg-duesseldorf.nrw.de/infos/Zivilprozess-der-Zukunft/Muenchener-Thesen-zum-Zivilprozess-der-Zukunft.pdf (Abruf am 26.02.2026, 18:27 Uhr).

Reformkommission Zivilprozess der Zukunft, Abschlussbericht der Reformkommission Zivilprozess der Zukunft im Auftrag des 3. Digitalgipfels der Justizministerinnen und Justizminister des Bundes und der Länder, abrufbar unter https://www.bmjv.de/SharedDocs/Publikationen/DE/Fachpublikationen/2025_Abschlussbericht_Reformkomission_Zivilprozess.html (Abruf am 20.04.2026, 21:09 Uhr).

Statistisches Bundesamt (Destatis), Statistischer Bericht – Zivilgerichte – 2025 (EVAS-Nummer 24231), 2026, abrufbar unter https://www.destatis.de/DE/Themen/Staat/Justiz-Rechtspflege/Publikationen/Downloads-Gerichte/statistischer-bericht-zivilgerichte-2100210257005.html (Abruf am 25.10.2026, 14:36 Uhr).

Turbék, Zoltán, Justiz und KI – Empfehlungen für einen verantwortungsvollen KI-Einsatz durch Gerichte, Friedrich-Naumann-Stiftung für die Freiheit (Hrsg.), Potsdam-Babelsberg 2026, abrufbar unter https://shop.freiheit.org/#!/Publikation/2184 (Abruf am 09.10.2026, 15:40 Uhr).

UNESCO, Guidelines for the Use of AI Systems in Courts and Tribunals, Paris 2025, abrufbar unter https://doi.org/10.58338/LIEY8089 (Abruf am 04.11.2026, 20:45 Uhr).

IX. Rechtsprechung

BGH, Urt. v. 6.10.2011 – RiZ (R) 7/10, MMR 2012, 128.

BVerfG, Urt. v. 14.7.1998 – 1 BvR 1640/97, BVerfGE 98, 218.

BVerfG, Urt. v. 24.1.2001 – 1 BvR 2623/9, 1 BvR 622/99, BVerfGE 103, 44.

BVerwG, Urteil vom 26.02.1997 - 6 C 3/96 (Lüneburg), NJW 1997, 2694.

X. Amtliche Dokumente und Rechtsvorschriften

BT-Drucks. 7/2459, Gesetzentwurf der Fraktion der CDU/CSU vom 07.08.1974, Entwurf eines Gesetzes zur Einführung der Zulassungsrevision gegen Berufungsurteile der Landgerichte in Zivilsachen.

BT-Drucks. 14/4722, Gesetzentwurf der Bundesregierung vom 24.11.2000, Entwurf eines Gesetzes zur Reform des Zivilprozesses.

BT-Drucks. 20/6520, Gesetzentwurf der Bundesregierung vom 24.04.2023, Entwurf eines Gesetzes zur Umsetzung der Richtlinie (EU) 2020/1828 über Verbandsklagen zum Schutz der Kollektivinteressen der Verbraucher und zur Aufhebung der Richtlinie 2009/22/EG (Verbandsklagenrichtlinienumsetzungsgesetz – VRUG)..

Gesetz zur Reform des Zivilprozesses, vom 27.07.2001, BGBl. I 2001, Nr. 40 vom 02.08.2001, S. 1887.

Justizaktenaufbewahrungsgesetz, vom 22. März 2005 (BGBl. I S. 837, 852), geändert durch Artikel 3 des Gesetzes vom 16. Oktober 2020 (BGBl. I S. 2187).

Justizaktenaufbewahrungsverordnung, vom 8. November 2021 (BGBl. I S. 4834), zuletzt geändert durch Artikel 3 des Gesetzes vom 29. März 2026 (BGBl. 2026 I Nr. 83).

Justizzuständigkeitsverordnung (Nordrhein-Westfalen), vom 4. Dezember 2024 (GV. NRW. S. 1144), zuletzt geändert durch Verordnung vom 5. Dezember 2025 (GV. NRW. S. 1152).

Richtlinie (EU) 2016/2102, über den barrierefreien Zugang zu den Websites und mobilen Anwendungen öffentlicher Stellen.

Richtlinie (EU) 2019/882, des Europäischen Parlaments und des Rates über die Barrierefreiheitsanforderungen für Produkte und Dienstleistungen.

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, 18.11.2024.

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. L, 2024/1689, 12.7.2024.

Verordnung (EU) 2026/1744, des Europäischen Parlaments und des Rates vom 8. Juli 2026 zur Änderung der Verordnungen (EU) 2024/1689, (EU) 2018/1139 und (EU) 2023/1230 im Hinblick auf die Vereinfachung der Umsetzung harmonisierter Vorschriften für künstliche Intelligenz (Digital-Omnibus-Verordnung zur KI), ABl. L, 2026/1744, 24.7.2026.