

KI in Klinik und Unternehmen

Nutzen belegen, Risiken begrenzen, Verantwortung sichern

Ein wissenschaftlich fundierter Leitfaden für Ethikbeauftragte, pflegerische Leitung und Unternehmensverantwortliche
Recherche- und Rechtsstand: 24. September 2026 · Deutschland und Europäische Union

Dokumenttyp: evidenzinformierter Fachbeitrag mit Praxisanhang. Keine systematische Übersichtsarbeit, keine konsentrierte medizinische Leitlinie und keine einzelfallbezogene Rechtsberatung.

Die vorgeschlagenen Organisationsverfahren sind eine begründete redaktionelle Synthese, kein wissenschaftlich validiertes Gesamtverfahren.

Steckbrief

Merkmal	Wert
Dokumenttyp	Evidenzinformierter Fachbeitrag mit Praxisanhang
Zielgruppe	Ethikbeauftragte, pflegerische Leitung, Unternehmensverantwortliche
Recherche- und Rechtsstand	24. September 2026
Umfang	7.975 Wörter, 24 Seiten
Belege	37 Literatureinträge, 70 Quellenadressen
Gliederung	11 nummerierte Abschnitte, 7 Tabellen

Inhalt

Kurzfassung	5
1. Die falsche Ausgangsfrage lautet: „Ist KI gut oder schlecht?“	5
2. Was die Forschung trägt – und was nicht	5
2.1 Klinischer Nutzen: ein konkreter Nachweis ist möglich	6
2.2 Dokumentationsassistent: Entlastung ist produkt- und endpunktspezifisch	6
2.3 Ein besseres Modell ist noch keine bessere Versorgung	6
2.4 Menschliche Beteiligung ist keine automatische Sicherheitsgarantie	7
2.5 Ein präziser Zielwert kann ethisch der falsche sein	7
2.6 In Unternehmen gibt es reale Gewinne – aber keine universelle Produktivitätsgarantie	7
3. Die Gefahren offen benennen – ohne unterschiedliche Evidenzlagen zu vermischen	7
Überzeugende Fehler und unbemerkte Auslassungen	7
Diskriminierung ist mehr als ein unvollständiger Datensatz	8
Datenschutz und Sicherheit sind eigenständige Prüfblöcke	8
Beziehung, Kompetenz und Arbeitsbedingungen können unter Druck geraten	8
Abhängigkeit, Konzentration und Ressourcenverbrauch gehören zur Ethikprüfung	9
Schwerwiegende Zukunftsrisiken: weder verschweigen noch als Gewissheit verkaufen	9
4. Werte bieten Orientierung; sie ersetzen keine Abwägung	9
5. Rechtsstand: Was gilt, was später greift und was nur Orientierung ist	9
5.1 Der KI-Rechtsrahmen hat mehrere Zeitachsen	9
5.2 Datenschutz und Arbeitsrecht warten nicht auf spätere Hochrisikofristen	11
5.3 Produktkonformität, fachlicher Standard und Ethik sind nicht dasselbe	11
5.4 Zuständigkeiten nicht verwechseln	11
6. Wer verantwortlich sein sollte – und warum die Pflegeleitung nicht fehlen darf	11
7. Von Prinzipien zu Entscheidungen: ein Verfahren mit sechs Prüfstufen	12
Stufe 1: Auftrag, Vergleich und Grenzen festlegen	13
Stufe 2: Belege statt Vorführungen verlangen	13
Stufe 3: Rechte, Betroffene und Zielkonflikte prüfen	13
Stufe 4: Einen begrenzten Pilot mit vorab festgelegten Kriterien planen	13
Stufe 5: Eine begrenzte, begründete Freigabe erteilen	14
Stufe 6: Betrieb, Änderungen und Ausstieg organisieren	14
8. Was sinnvoll gemessen werden sollte	14
Warum „95 Prozent Genauigkeit“ wenig sagen kann	14
9. Drei illustrative Anwendungssituationen	15
Krankenhaus: Ein Frühwarnsystem erzeugt mehr Hinweise als das Team bearbeiten kann	15
Krankenhaus: Ein Assistent hört das Behandlungsgespräch mit	15
Konzern: Ein System priorisiert Bewerbungen	15
10. Wo und wie Ethikbeauftragte sich informieren können	16
11. Was Gesellschaft und Organisationen als Lösung anstreben sollten	17
A. Die ersten 90 Tage: eine realistische Arbeitsstruktur aufbauen	17

B. Zwölf Fragen, die vor einer Beschaffung beantwortet sein sollten	18
C. Inhalt einer nachvollziehbaren Entscheidungsvorlage	18
D. Wann eine Anwendung nicht unverändert weiterlaufen sollte	19
E. Kleines Glossar	19
Methodik, Grenzen und Hinweise vor Veröffentlichung	19
Rechercheansatz	19
Bewusst getrennte Aussageebenen	20
Redaktioneller Prüfstand	20
Literatur und Primärquellen	20
Empirische Forschung	20
Ethik, Organisation, Sicherheit und Evaluationsmethodik	21
Recht und regulatorische Orientierung	22
12 Rechtliche Hinweise	23
12.1 Keine Beratung	23
12.2 Keine Zusicherung, keine Konformitätsaussage	23
12.3 Rechtsstand	23
12.4 Haftung	23
12.5 Verweise auf fremde Inhalte	23
12.6 Marken und Produktbezeichnungen	23
12.7 Urheberrecht und Nutzung	23
12.8 Keine personenbezogenen Daten	24
12.9 Unabhängigkeit	24
12.10 Anwendbares Recht	24

Impressum und Verantwortlichkeit

Die folgenden Angaben sind wörtlich dem Impressum von fluxlane.eu entnommen, abgerufen am 24. September 2026. Sie wurden nicht ergänzt und nicht umformuliert.

Angabe	Inhalt
Angaben gemäß § 5 DDG	WebTrades GmbH Am Ölberg 2 55459 Grolsheim Deutschland
Vertreten durch	Geschäftsführerin: Laurice Bussmer
Kontakt	Telefon: 0170 890 78 00 E-Mail: info@fluxlane.eu
Registereintrag	Eintragung im Handelsregister Registergericht: Amtsgericht Mainz Registernummer: HRB 54479
Umsatzsteuer-ID	Umsatzsteuer-Identifikationsnummer gemäß § 27 a Umsatzsteuergesetz: DE461650292
Redaktionell verantwortlich	Verantwortlich für den Inhalt nach § 18 Abs. 2 MStV: Laurice Bussmer, Am Ölberg 2, 55459 Grolsheim
Verbraucherstreitbeilegung	Wir sind nicht bereit oder verpflichtet, an Streitbeilegungsverfahren vor einer Verbraucherschlichtungsstelle teilzunehmen.

Redaktioneller Hinweis des Herausgebers (wörtlich)

„fluxlane ist ein unabhängiges, redaktionelles Fachportal für Robotik, autonome physische Systeme und verkörperte KI. Der redaktionelle Geltungsbereich umfasst Industrie und Produktion, Logistik, Medizin, Pflege und Rehabilitation, Service und Haushalt sowie Forschung, Hersteller, Markt und Regulierung; künstliche Intelligenz gehört dazu, auch ohne physischen Körper. Die Inhalte dienen der allgemeinen Information und stellen keine Rechts-, Kauf- oder Anlageberatung dar. Angaben zu Preisen, Verfügbarkeit und Terminen erfolgen ohne Gewähr und können sich jederzeit ändern.“

Zitierweise

FluxLane (2026): *KI in Klinik und Unternehmen – Nutzen belegen, Risiken begrenzen, Verantwortung sichern*. Fassung 1.0, Stand 24. September 2026. Herausgegeben von der WebTrades GmbH, Grolsheim, fluxlane.eu.

Herstellung dieses Dokuments

Das vorliegende PDF ist maschinell aus der Markdown-Fassung gesetzt. Vor dem Satz prüft das Erzeugungsskript Belegziffern gegen das Literaturverzeichnis, die Anrede und benannte Datumsfallen; schlägt eine Prüfung fehl, entsteht kein Dokument. Gesetzt wurde mit IBM Plex Sans und Sora.

Abbildungsnachweis

Dieses Dokument enthält keine Fotografien. Alle Abbildungen sind Diagramme, die beim Satz aus den im Text belegten Werten gezeichnet wurden; die jeweilige Fundstelle steht unter der Abbildung. Die Diagramme sind KI-gestützt erstellt.

Kurzfassung

Künstliche Intelligenz kann Versorgung und Arbeit verbessern. Sie kann aber auch überzeugend falsche Informationen liefern, bestehende Benachteiligungen verstärken, Überwachung ausweiten und Verantwortung so verteilen, dass im Schadensfall niemand wirksam zuständig ist. Die wissenschaftlich angemessene Antwort besteht weder in pauschalem Vertrauen noch in pauschaler Ablehnung. Entscheidend sind die konkrete Aufgabe, die Qualität der Vergleichspraxis, die betroffenen Menschen und die Bedingungen des Einsatzes.

Studien zeigen bereits messbaren Nutzen in bestimmten Anwendungen, darunter KI-unterstützte Mammographie und einzelne Dokumentations- und Arbeitsassistenzsysteme. Zugleich belegen Untersuchungen, dass gute Modellleistungen keine guten Mensch-KI-Entscheidungen garantieren. Sicherheits- und Nutzenbehauptungen müssen deshalb am vollständigen Arbeitsablauf geprüft werden, nicht nur am Modell.

Für Ethikbeauftragte folgt daraus ein anspruchsvolles, aber begrenztes Mandat: Sie sollen Werte- und Interessenkonflikte sichtbar machen, Betroffenen Gehör verschaffen und nachvollziehbare Entscheidungen ermöglichen. Sie ersetzen weder die verantwortliche Leitung noch Datenschutz, IT-Sicherheit, Medizinproduktekompetenz oder die fachliche Verantwortung der Anwender. Für Krankenhäuser empfehlen wir ausdrücklich die Beteiligung der pflegerischen Leitung an der Gestaltung: Sie kann beurteilen, ob ein versprochener Nutzen unter realen Personal-, Schicht- und Versorgungsbedingungen überhaupt erreichbar ist.

Der Kern dieses Beitrags ist ein Vorschlag für risikoadäquate Entscheidungen: Zweck und Alternativen klären, Evidenz prüfen, Betroffene beteiligen, begrenzt erproben, begründet freigeben und fortlaufend überwachen. Dazu gehören vorab festgelegte Abbruchkriterien, wirksame Beschwerdewege und eine sichere Rückkehr zu einem Ersatzverfahren. Rechtliche Mindestanforderungen, fachliche Empfehlungen und eigene Umsetzungsvorschläge werden ausdrücklich getrennt.

1. Die falsche Ausgangsfrage lautet: „Ist KI gut oder schlecht?“

Ein System, das öffentliche Texte zusammenfasst, ist ethisch anders zu beurteilen als eines, das Bewerber vorsortiert, eine Sepsiswarnung auslöst oder einen Roboter im unmittelbaren Kontakt mit pflegebedürftigen Menschen steuert. Selbst dasselbe Sprachmodell kann in einem Kontext hilfreich und in einem anderen unververtretbar sein. Im Experiment mit Unternehmensberatern waren Leistungsgewinne und Leistungseinbußen tatsächlich innerhalb desselben untersuchten Werkzeugtyps zu beobachten.[11]

Gegenstand einer verantwortlichen Prüfung ist deshalb nicht „die KI“, sondern eine konkrete Kombination: Modell und Version, verwendete Daten, Benutzeroberfläche, Zugriffsrechte, Arbeitsablauf, Nutzergruppe, betroffene Population und institutionelle Anreize. Ein fehlerfreier Hinweis nützt wenig, wenn nachts niemand darauf reagieren kann. Eine durchschnittlich schnellere Dokumentation verbessert die Versorgung nicht automatisch, wenn die gewonnene Zeit vollständig in zusätzliche Fallzahlen umgewandelt wird. Diese beiden Beispiele sind organisatorische Überlegungen, keine allgemeinen empirischen Befunde.

Ebenso wichtig ist der Vergleichsmaßstab. Menschliche Routine ist nicht fehlerfrei. Ein System sollte nicht mit einer idealisierten menschlichen Leistung verglichen werden, sondern mit der tatsächlich verfügbaren Versorgung oder Arbeitsweise und mit realistischen Alternativen. Dazu können bessere Personalausstattung, vereinfachte Formulare, konventionelle Software oder ein veränderter Prozess gehören. Die Unterlassung einer wirksamen Verbesserung kann ebenfalls Schaden verursachen. Daraus ergibt sich aus ethischer Sicht kein Innovationsgebot um jeden Preis, sondern die Forderung nach einer begründeten Abwägung.

Die leitende Frage lautet: **Verbessert dieser konkrete Einsatz eine für Menschen relevante Situation – und sind seine verbleibenden Risiken, Verteilungswirkungen und Eingriffe vertretbar?**

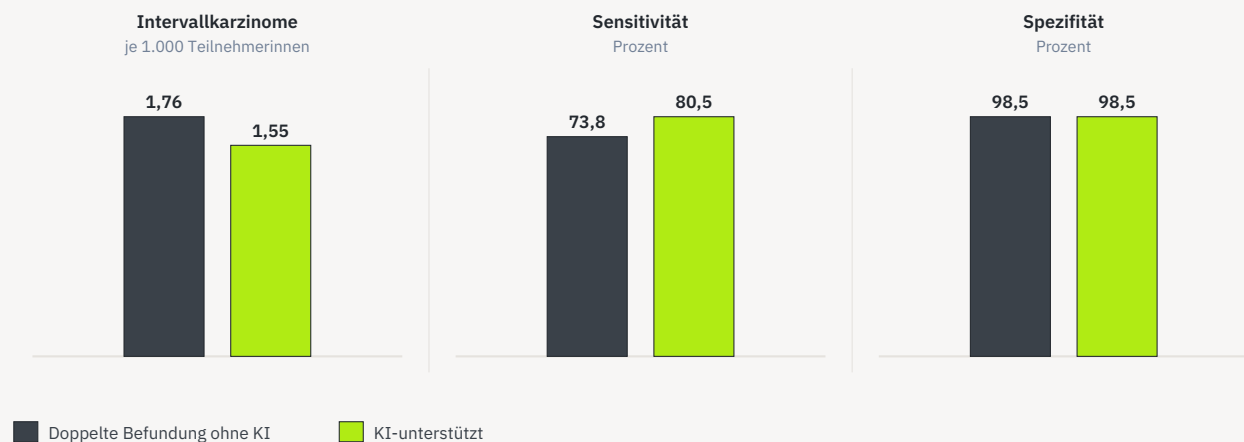
2. Was die Forschung trägt – und was nicht

Die folgenden Studien wurden gezielt ausgewählt, weil sie unterschiedliche Erkenntnisprobleme sichtbar machen. Sie bilden kein repräsentatives Gesamtbild aller KI-Anwendungen. Produktversionen, Aufgaben, Populationen und Endpunkte unterscheiden sich; ihre Resultate dürfen nicht zu einer gemeinsamen „KI-Erfolgsquote“

zusammengerechnet werden.

2.1 Klinischer Nutzen: ein konkreter Nachweis ist möglich

Abbildung 2 – MASAI-Studie: gemessene Unterschiede



Werte aus Abschnitt 2.1 dieses Leitfadens, Beleg [3]. Jedes Paneel hat eine eigene Skala, weil die drei Größen verschiedene Einheiten tragen. Die Intervallkarzinomrate erfüllte die vorab definierte Nichtunterlegenheitsgrenze bei einem Verhältnis von 0,88 (95%-KI 0,65–1,18); das ist kein Nachweis einer gesicherten Reduktion und keiner geringeren Sterblichkeit.

Die schwedische MASAI-Studie randomisierte 105.934 Frauen zu KI-unterstütztem Mammographie-Screening oder standardmäßiger doppelter menschlicher Befundung; 19 wurden aus der Analyse ausgeschlossen. Im KI-Arm unterstützte das System die Befundung und die Zuordnung zu einfacher beziehungsweise doppelter menschlicher Befundung. Die 2026 veröffentlichte Analyse berichtete 1,55 gegenüber 1,76 Intervallkarzinomen je 1.000 Teilnehmerinnen. Das Verhältnis betrug 0,88, mit einem 95%-Konfidenzintervall von 0,65 bis 1,18. Die vorab definierte Nichtunterlegenheitsgrenze wurde eingehalten. Die Sensitivität lag bei 80,5 gegenüber 73,8 Prozent, die Spezifität in beiden Gruppen bei 98,5 Prozent.[3]

Das ist belastbare Evidenz für den untersuchten Screeningablauf. Es ist aber weder der Nachweis einer statistisch gesicherten Reduktion von Intervallkarzinomen um zwölf Prozent noch einer geringeren Brustkrebssterblichkeit. Nichtunterlegenheit innerhalb einer gewählten Grenze bedeutet nicht Gleichheit in jeder Hinsicht. Die Übertragbarkeit auf andere Produkte, Populationen und Versorgungssysteme bleibt eine eigene Prüffrage.

2.2 Dokumentationsassistent: Entlastung ist produkt- und endpunktspezifisch

Eine pragmatische randomisierte Studie mit 238 ambulant tätigen Ärzten aus 14 Fachrichtungen verglich über zwei Monate zwei KI-Dokumentationsassistenten mit der üblichen Dokumentation. Für Nabla sank die im Dokumentationsfenster verbrachte Zeit gegenüber der Kontrollgruppe um 9,5 Prozent; das 95%-Konfidenzintervall lag zwischen –17,2 und –1,8 Prozent. Für DAX betrug die Veränderung –1,7 Prozent, mit einem Intervall von –9,4 bis +5,9 Prozent; ein entsprechender Effekt war damit nicht statistisch nachgewiesen.[7]

Gemessen wurde nicht die gesamte Arbeitszeit. Die Ergebnisse belegen auch nicht, dass jede Dokumentations-KI die gleiche Entlastung bringt. Für eine lokale Einführung empfehlen sich deshalb Messungen von Korrekturaufwand, Dokumentationsqualität und tatsächlicher Nettoentlastung. Das sind hier abgeleitete Prüfvorschläge, keine zusätzlichen Ergebnisse der Studie.

2.3 Ein besseres Modell ist noch keine bessere Versorgung

Eine externe Untersuchung des damaligen Epic-Sepsismodells ergab 2021 an einer US-Einrichtung eine AUROC von 0,63. Bei der untersuchten Alarmschwelle wurden 67 Prozent der Sepsisfälle nicht erkannt; bei 18 Prozent der stationären Aufenthalte wäre ein Alarm ausgelöst worden. Das war eine retrospektive Validierung, kein Nachweis, dass das System diese Patienten tatsächlich geschädigt hatte.[5]

Die historische Kritik darf nicht ungeprüft auf die aktuelle Produktgeneration übertragen werden. Eine 2026 veröffentlichte prospektive externe Untersuchung von Version 2 an vier US-Gesundheitssystemen mit 227.091 Aufenthalten berichtete auf Aufenthaltsebene AUROC-Werte von 0,82 bis 0,92. Bei standortspezifischen Schwellen

für 60 Prozent Sensitivität lag der positive Vorhersagewert jedoch nur zwischen 13 und 26 Prozent. An allen vier Standorten war das Modell zuvor lokal angepasst worden; die Ergebnisse beschreiben nicht die Leistung des unveränderten Basismodells.[6]

Die gemeinsame Lehre: Versionen und Einsatzorte zählen. Verbesserte prognostische Leistung ist erfreulich, ersetzt aber keinen Nachweis besserer patientenrelevanter Ergebnisse. Entscheidend bleiben Fehlalarme, Reaktionsmöglichkeiten und der Nutzen des gesamten Versorgungspfads.

2.4 Menschliche Beteiligung ist keine automatische Sicherheitsgarantie

In einer randomisierten Studie mit 50 Ärzten verbesserte der Zugang zu GPT-4 den diagnostischen Begründungsscore bei klinischen Fallvignetten nicht statistisch nachweisbar: Die adjustierte Differenz betrug zwei Prozentpunkte, mit einem 95%-Konfidenzintervall von -4 bis +8. Das Experiment untersuchte begrenzte schriftliche Fälle, keine Behandlung realer Patienten.[4]

Ein anderes Experiment mit 27 Radiologen zeigte, dass absichtlich falsche, als KI-Empfehlungen ausgegebene BI-RADS-Einstufungen die Befundung in allen untersuchten Erfahrungsgruppen beeinflussten. Diese Versuchsanordnung belegt eine Anfälligkeit für Automationsgläubigkeit, nicht die Fehlerhäufigkeit eines real eingesetzten Produkts.[8]

Daraus folgt eine praktische Empfehlung: Die Formel „Ein Mensch schaut darüber“ reicht als Sicherheitsnachweis nicht aus. Es muss geprüft werden, ob dieser Mensch einen Fehler unter realen Bedingungen erkennen und korrigieren kann. Umgekehrt rechtfertigen die Studien nicht, Menschen pauschal aus folgenreichen Entscheidungen zu entfernen. Aufgabenleistung, Verantwortlichkeit und die Rechte Betroffener sind unterschiedliche Bewertungsdimensionen.

2.5 Ein präziser Zielwert kann ethisch der falsche sein

Obermeyer und Kollegen untersuchten einen tatsächlich verwendeten US-Algorithmus zur Steuerung zusätzlicher Versorgung. Er nutzte erwartete Gesundheitsausgaben als Stellvertreter für gesundheitlichen Bedarf. Bei gleichem Risikoscore waren Schwarze Patienten kränker als weiße. Der Mechanismus lag wesentlich darin, dass Ausgaben wegen ungleichen Versorgungszugangs den Bedarf nicht für alle Gruppen gleich abbildeten.[9]

Die Übertragung auf Deutschland ist nicht automatisch möglich. Die grundsätzliche Prüffrage ist jedoch unmittelbar relevant: Was optimiert ein System tatsächlich? Kosten, Umsatz, frühere Beurteilungen und vergangene Nutzung sind nicht zwangsläufig gute Stellvertreter für Gesundheit, Eignung oder Unterstützungsbedarf. Ein Modell kann sein statistisches Ziel zuverlässig erreichen und damit dennoch eine ungerechte Entscheidung unterstützen.

2.6 In Unternehmen gibt es reale Gewinne – aber keine universelle Produktivitätsgarantie

Die finale Veröffentlichung von „Generative AI at Work“ untersuchte die gestaffelte Einführung eines Assistenten bei 5.172 Kundendienstbeschäftigten. Im Mittel wurden etwa 15 Prozent mehr Anliegen pro Stunde gelöst. Unerfahrenere Beschäftigte profitierten stärker. Das war eine Feldstudie in einem Unternehmen, keine randomisierte Aussage über die gesamte Wirtschaft oder langfristige Beschäftigungseffekte.[10]

Ein präregistriertes Experiment mit 758 Unternehmensberatern zeigte die Abhängigkeit von der Aufgabe besonders deutlich: Bei ausgewählten Aufgaben innerhalb der untersuchten KI-Fähigkeiten wurden 12,2 Prozent mehr Aufgaben und diese im Mittel 25,1 Prozent schneller erledigt. Bei einer anderen, komplexen Managementaufgabe waren dagegen 84,5 Prozent der Lösungen ohne KI richtig, gegenüber 60 beziehungsweise 70,6 Prozent in den beiden KI-Gruppen.[11]

Eine verantwortliche Organisation sollte daraus weder eine pauschale Einführungspflicht noch ein pauschales Verbot ableiten. Sie sollte Freigaben auf Aufgaben zuschneiden und Qualität ebenso ernst nehmen wie Geschwindigkeit.

3. Die Gefahren offen benennen – ohne unterschiedliche Evidenzlagen zu vermischen

Überzeugende Fehler und unbemerkte Auslassungen

Generative Systeme können sachlich falsche, unbelegte oder widersprüchliche Antworten formulieren. Ein sprachlich überzeugender Text ist kein Wahrheitsnachweis. Auch ausgelassene Informationen können entscheidend sein: etwa eine Kontraindikation in einer Zusammenfassung oder eine Ausnahme in einer internen Regel. Die WHO behandelt solche Risiken ausdrücklich in ihrer Orientierung zu großen multimodalen Modellen.[2]

Unsere Empfehlung lautet: In folgenreichen Abläufen müssen kritische Aussagen anhand geeigneter Originalinformationen überprüfbar sein. Eine Quellenangabe allein genügt nicht; die Quelle muss existieren, zur Aussage passen und den jeweiligen Fall tragen. Ein System mit Dokumentensuche, sogenanntem Retrieval-Augmented Generation oder RAG, kann die Wissensgrundlage verbessern. Es beseitigt Fehler aber nicht automatisch. Die Datenschutzkonferenz behandelt zudem spezifische Zugriffs- und Datenschutzfragen solcher Architekturen.[19]

Diskriminierung ist mehr als ein unvollständiger Datensatz

Prüfgegenstand sind Daten, Zielvariable, Entscheidungsschwelle und Folgen. Gleiche durchschnittliche Genauigkeit schließt unterschiedliche Fehlerraten oder Belastungen nicht aus. Besonders schwer sichtbar bleiben kleine oder in den Daten kaum erfasste Gruppen. Die STANDING-Together-Empfehlungen richten den Blick auf Datenzusammensetzung, Grenzen der Repräsentation und die Prüfung von Benachteiligungen.[30]

Als Umsetzung schlagen wir vor: Relevante Gruppen vorab begründen, Leistungs- und Ergebnisunterschiede mit Unsicherheiten ausweisen und Betroffene an der Auswahl der Kriterien beteiligen. Fehlende Daten sind dabei eine Erkenntnisgrenze, kein Beweis gleicher Behandlung. Die Erhebung sensibler Merkmale zur Fairnessprüfung benötigt ihrerseits eine rechtlich tragfähige Gestaltung. Gerechtigkeit lässt sich außerdem nicht auf eine universell richtige Kennzahl reduzieren; Konflikte zwischen Kriterien müssen begründet entschieden werden.

Datenschutz und Sicherheit sind eigenständige Prüfblöcke

Ein nützliches System kann unzulässig auf Daten zugreifen. Ein zulässiger Datenfluss kann technisch unsicher sein. Bei generativen Assistenzsystemen kommen Risiken durch manipulierte Eingaben oder Dokumente hinzu: Fremde Inhalte können als Anweisungen missverstanden werden. Besonders folgenreich wird dies, wenn ein Assistent zusätzlich Nachrichten versenden, Daten ändern oder externe Werkzeuge bedienen darf.[17,34]

Unsere Empfehlungen sind abgestufte Rechte, eine Trennung von Informationszugriff und Handlungsvollmacht, überprüfbare Datenflüsse und zusätzliche Freigaben bei folgenreichen Aktionen. Vertrauliche Daten gehören nicht allein deshalb in ein System, weil dessen Oberfläche bequem erreichbar ist. Umgekehrt macht ein Betrieb im eigenen Rechenzentrum eine Verarbeitung nicht automatisch rechtmäßig. Zweck, Berechtigungen und Rechtsgrundlage bleiben zu prüfen.[18,19,22]

Beziehung, Kompetenz und Arbeitsbedingungen können unter Druck geraten

Der ICN-Ethikkodex fordert für KI und entsprechende Geräte ausdrücklich personenzentrierte Pflege und den Erhalt menschlicher Beziehungen.[12] Daraus folgt nicht, dass jede technische Interaktion entmenslicht. Eine gut gestaltete Assistenz kann Zeit für Gespräche eröffnen oder Selbstständigkeit stärken. Ob das tatsächlich geschieht, hängt von Einsatz und Organisation ab.

Als prüfbare Umsetzung empfehlen wir, erreichbare menschliche Ansprechpartner, erlebte Selbstbestimmung und unerfüllten Unterstützungsbedarf zu erfassen. Ein Gesprächsroboter sollte nicht verschleiern, dass er ein technisches System ist. Bei verletzlichen Personen sind Verständnis, Freiwilligkeit und Ausweichmöglichkeiten besonders sorgfältig zu betrachten. Befürchtungen zu Kompetenzverlust oder dauerhafter Beziehungssubstitution sind ernst zu nehmen; die hier ausgewählten Studien quantifizieren deren langfristiges Ausmaß jedoch nicht.

Für körpernah eingesetzte Robotik empfehlen wir zusätzlich eine eigenständige Prüfung physischer Gefährdungen und der sicheren Übergabe an Menschen: etwa bei Kollision, unerwarteter Bewegung, Ausfall oder einer nicht verstandenen Aufforderung. Je nach Aufgabe gehören Hygiene, Zugänglichkeit und die Möglichkeit, sich einer Interaktion zu entziehen, dazu. Die in diesem Beitrag ausgewählten Softwarestudien liefern keinen Wirksamkeits- oder Sicherheitsnachweis für einen Pflegeroboter. Seine Bewertung benötigt zusätzliche produktspezifische Evidenz und Fachkompetenz.

Abhängigkeit, Konzentration und Ressourcenverbrauch gehören zur Ethikprüfung

Die Bindung an wenige Anbieter kann Wechselkosten und institutionelle Abhängigkeit schaffen. Ein späterer Modellwechsel kann Leistung, Datenverarbeitung und Verfügbarkeit verändern. Auch Energie- und Infrastrukturbedarf sind relevant. Die IEA untersucht hierfür die Entwicklung von Rechenzentren und KI; Prognosen sind jedoch keine bereits eingetretenen Verbräuche, und Rechenzentrumsstrom darf nicht vollständig der KI zugerechnet werden.[33]

Unsere Empfehlung ist eine aufgabenbezogene Ressourcenbilanz: Welcher Nutzen rechtfertigt welchen Rechen-, Prüf- und Personalaufwand? Ist ein kleineres Modell oder eine konventionelle Lösung ausreichend? Anbieterwechsel und Ausfall sollten vor der Beschaffung praktisch durchgespielt werden. Dabei sind wirtschaftliche Effizienz, Versorgungssicherheit und Nachhaltigkeit gemeinsam zu betrachten, nicht gegeneinander auszuspielen.

Schwerwiegende Zukunftsrisiken: weder verschweigen noch als Gewissheit verkaufen

Der International AI Safety Report 2026 unterscheidet Risiken durch schädliche Nutzung, Fehlfunktionen und systemische Entwicklungen. Er behandelt auch mögliche Kontrollverluste bei leistungsfähigeren allgemeinen Systemen. Solche Szenarien sind nicht mit der heute belegten Fehlerquote eines klinischen Assistenten gleichzusetzen. Eintrittswahrscheinlichkeiten und Übertragbarkeit experimenteller Befunde bleiben vielfach unsicher.[32]

Für Einrichtungen empfehlen wir eine proportionale Konsequenz: Mit wachsender Autonomie, Systemanbindung und möglichem Schadensumfang steigen die Anforderungen an Begrenzung, unabhängige Prüfung und Rückholbarkeit. Für Gesellschaft und Politik sind gemeinsame Sicherheitsforschung, Vorfallauswertung und grenzüberschreitende Regeln sinnvoll. Eine einzelne Ethikbeauftragte kann diese Aufgaben nicht stellvertretend für Staaten oder Hersteller erfüllen.

4. Werte bieten Orientierung; sie ersetzen keine Abwägung

Die WHO nennt für KI im Gesundheitswesen unter anderem Autonomie, Wohlergehen und Sicherheit, Transparenz, Verantwortlichkeit, Gerechtigkeit und Nachhaltigkeit.[1] Diese Orientierung ist auch außerhalb des Krankenhauses hilfreich. Sie beantwortet aber nicht automatisch, wie ein konkreter Konflikt gelöst werden soll.

Mehr Sensitivität kann mehr Fehlalarme bedeuten. Personalisierung kann zusätzliche Datenverarbeitung verlangen. Eine leicht erklärbare Methode kann in einem bestimmten Fall weniger leistungsfähig sein. Umgekehrt macht eine komplizierte Methode eine Entscheidung nicht grundsätzlich unvertretbar. Entscheidend ist, ob die Gründe, Grenzen und Folgen für den jeweiligen Zweck ausreichend prüfbar sind.

Unser Vorschlag ist, in jedem wesentlichen Entscheidungsvorgang vier Fragen ausdrücklich zu beantworten: Welcher menschliche Nutzen wird angestrebt? Welche Rechte und Belastungen sind betroffen? Welche Alternative ist verfügbar? Wer darf die verbleibende Unsicherheit unter welchen Bedingungen verantworten? Das Ergebnis sollte eine Begründung enthalten, nicht nur ein grünes Ampelfeld.

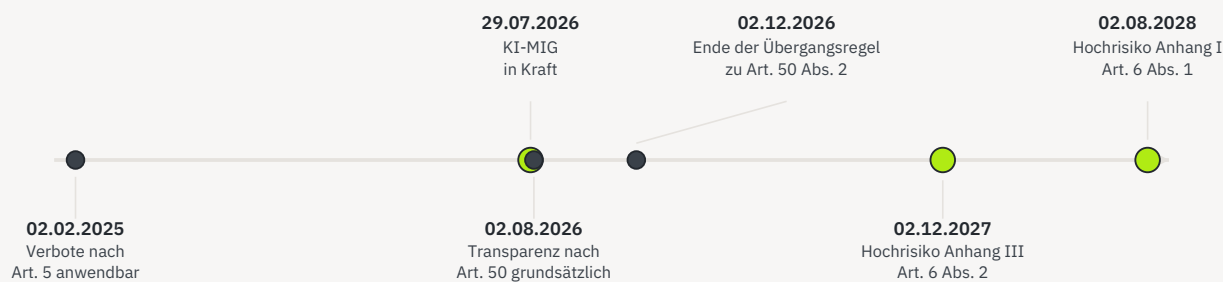
Rechtskonformität ist notwendig, aber kein vollständiger Nachweis ethischer Vertretbarkeit. Umgekehrt macht eine gute ethische Absicht eine rechtswidrige Verarbeitung nicht zulässig.

5. Rechtsstand: Was gilt, was später greift und was nur Orientierung ist

Dieser Abschnitt betrifft Deutschland und die EU. Er ist ein Einstieg in die Prüfung, keine vollständige Behandlung aller Haftungs-, Berufs-, Sozial-, Forschungs- und sektoralen Vorschriften. Für Konzernstandorte außerhalb der EU sind weitere Rechtsordnungen zu berücksichtigen.

5.1 Der KI-Rechtsrahmen hat mehrere Zeitachsen

Abbildung 1 – Geltungstermine des KI-Rechtsrahmens



Werte aus Abschnitt 5.1 und 5.4 dieses Leitfadens. Rechtsgrundlagen: Verordnung (EU) 2024/1689 in der durch Verordnung (EU) 2026/1744 geänderten Fassung, Art. 113 Buchst. c; KI-MIG vom 22.07.2026. Stand 24.09.2026.

Der AI Act, Verordnung (EU) 2024/1689, wurde durch die Verordnung (EU) 2026/1744 geändert. Dieser „AI-Omnibus“ ist seit 27. Juli 2026 in Kraft. Ein älterer Fristenplan darf deshalb nicht unverändert verwendet werden.[20,21]

Regelungsbereich	Stand am 24. September 2026	Konsequenz für die Einrichtung
Bereits anwendbare Verbote nach Art. 5	Die seit Februar 2025 anwendbaren Verbote gelten weiter; neu eingefügte Verbote besitzen teilweise eigene spätere Termine.	Unzulässige Funktionen nicht als gewöhnliches Pilotprojekt behandeln.
KI-Kompetenz, Art. 4	Die geänderte Vorschrift verpflichtet Anbieter und Betreiber zu Maßnahmen zur Förderung der Kompetenzentwicklung, nicht zur Garantie eines bestimmten individuellen Kompetenzniveaus.	Auf Aufgaben und Risiken bezogene Qualifizierung organisieren; kein gesetzlich vorgeschriebenes allgemeines „KI-Führerschein“-Zertifikat behaupten.
Transparenz, Art. 50	Grundsätzlich seit 2. August 2026 anwendbar. Für Anbieter bestimmter vor diesem Datum vermarkteter generativer Systeme besteht eine begrenzte Übergangsregel zu Art. 50 Abs. 2 bis 2. Dezember 2026.	Nicht sämtliche Transparenzpflichten für verschoben halten; Adressat und konkrete Funktion prüfen.
Hochrisiko nach Art. 6 Abs. 2 / Anhang III	Kapitel III Abschnitte 1–3, mit Ausnahme von Art. 6 Abs. 5, ab 2. Dezember 2027.	Einstufung und Umsetzungsplanung auf den aktuellen Rechtsakt stützen.
Produktbezogenes Hochrisiko nach Art. 6 Abs. 1 / Anhang I	Die entsprechenden Vorschriften ab 2. August 2028.	Medizinprodukte- und andere Produktpflichten gelten unabhängig davon weiter.

Ein Krankenhaus ist nicht als Ganzes „Hochrisiko-KI“. Zu prüfen sind Zweck und Funktion des einzelnen Systems sowie die Voraussetzungen und Ausnahmen der jeweiligen Einstufung. Bei produktbezogener KI ist insbesondere die Verbindung mit den einschlägigen Produktvorschriften und einer erforderlichen Drittstellen-Konformitätsbewertung relevant. Der bloße Einsatz in einer Klinik macht allgemeine Software noch nicht zum Medizinprodukt.[20,21,23]

Vor der Bestimmung der Pflichten steht außerdem die Rollenprüfung: Eine Einrichtung kann Systeme nur betreiben, aber auch selbst entwickeln oder so verändern, dass zusätzliche Anbieterpflichten relevant werden. Eigenentwicklungen, Integration und wesentliche Veränderungen sind deshalb ausdrücklich zu erfassen; insbesondere Art. 25 ist im einschlägigen Anwendungsbereich und zum jeweiligen Anwendungszeitpunkt zu berücksichtigen.[20,21]

Für einschlägige Hochrisikosysteme verlangt Art. 26 nach dem jeweiligen Anwendungsbeginn unter anderem bestimmungsgemäße Nutzung, geeignete Aufsicht und Überwachung. Aufsichtspersonen benötigen Kompetenz, Schulung, Befugnisse und Unterstützung. Die organisatorische Empfehlung dieses Beitrags lautet, entsprechende Voraussetzungen bereits bei heutigen folgenreichen Einsätzen herzustellen. Damit wird die spätere AI-Act-Pflicht nicht fälschlich als heute schon generell anwendbar ausgegeben.[20,21]

Die nach dem einschlägigen Anwendungsbeginn zu erfüllende Pflicht zur Grundrechte-Folgenabschätzung nach Art. 27 trifft nicht unterschiedslos jeden Konzern oder jedes medizinische KI-Produkt. Sie betrifft bestimmte Betreiber bestimmter Anhang-III-Systeme, unter anderem öffentlich-rechtliche Stellen und private Erbringer öffentlicher

Dienste; gesetzliche Ausnahmen und besondere Finanzanwendungen sind zu beachten. Sie ist von der Datenschutz-Folgenabschätzung zu unterscheiden, auch wenn passende Bestandteile übernommen oder miteinander verknüpft werden können.[20,21]

5.2 Datenschutz und Arbeitsrecht warten nicht auf spätere Hochrisikofristen

Die DSGVO verlangt eine tragfähige Grundlage für personenbezogene Verarbeitung. Bei Gesundheitsdaten ist zusätzlich eine einschlägige Ausnahme nach Art. 9 erforderlich. Ein Auftragsverarbeitungsvertrag ist nicht selbst die Rechtsgrundlage. Bei voraussichtlich hohem Risiko ist nach Art. 35 eine Datenschutz-Folgenabschätzung erforderlich; nicht jede Verwendung des Etiketts „KI“ löst sie automatisch aus.[22]

Art. 22 betrifft ausschließlich automatisierte Entscheidungen mit rechtlicher oder ähnlich erheblicher Wirkung, vorbehaltlich seiner Ausnahmen und Schutzvorkehrungen. Nicht jede Assistenz fällt darunter. Eine nur formale menschliche Bestätigung darf jedoch nicht als substantielle Prüfung ausgegeben werden.[18,22]

Beschäftigtenvertretungen sind frühzeitig einzubeziehen, wo die einschlägigen Beteiligungsrechte greifen. § 90 BetrVG erfasst ausdrücklich die Planung von Arbeitsverfahren einschließlich KI; § 87 Abs. 1 Nr. 6 betrifft technische Einrichtungen zur Verhaltens- oder Leistungsüberwachung. Öffentliche und kirchliche Krankenhausträger unterliegen teilweise anderen Vertretungsregeln. Daneben bleibt das Antidiskriminierungsrecht relevant.[24,25]

Eine besonders deutliche rote Linie ist die grundsätzlich verbotene KI-gestützte Ableitung von Emotionen auf Grundlage biometrischer Daten am Arbeitsplatz nach Art. 5 Abs. 1 Buchst. f AI Act. Die vorgesehenen medizinischen oder sicherheitsbezogenen Ausnahmen sind zweckbezogen; der Standort „Krankenhaus“ schafft keine pauschale Ausnahme für Personalbewertung.[20]

5.3 Produktkonformität, fachlicher Standard und Ethik sind nicht dasselbe

Bei medizinischer Software sind Zweckbestimmung und mögliche Folgen wesentlich. Regel 11 der MDR differenziert nach Funktion und Risiko; „medizinische KI ist immer Klasse IIa“ wäre falsch. Konformität im jeweiligen Geltungsbereich ersetzt nicht die Prüfung, ob das System für die konkrete lokale Verwendung geeignet ist. Diese zweite Aussage ist eine fachlich-organisatorische Folgerung, keine Behauptung, eine CE-Kennzeichnung sei bedeutungslos.[23,36]

§ 630a Abs. 2 BGB knüpft medizinische Behandlung grundsätzlich an die allgemein anerkannten fachlichen Standards zum Behandlungszeitpunkt an. Eine Ethikempfehlung wird dadurch nicht automatisch zur verbindlichen Rechtsnorm. Die Frage, ob und wie sie für einen fachlichen Standard relevant ist, muss gesondert betrachtet werden.[27]

Auch WHO-, AEM- und NIST-Dokumente sind keine deutschen Gesetze. ISO/IEC 42001 betrifft ein KI-Managementsystem. Ein entsprechendes Zertifikat belegt nicht automatisch die klinische Wirksamkeit eines Produkts oder die Rechtmäßigkeit jedes Einsatzes. Dafür sind Zertifikatsumfang und anwendungsspezifische Nachweise zu prüfen.[31]

5.4 Zuständigkeiten nicht verwechseln

Das deutsche Gesetz zur Marktüberwachung und Innovationsförderung von künstlicher Intelligenz (KI-MIG) gilt seit 29. Juli 2026. Die Bundesnetzagentur besitzt zentrale und auffangende Aufgaben; daneben bleiben sektorale und länderspezifische Zuständigkeiten relevant. § 6 regelt eine zentrale Anlaufstelle. Daraus folgt weder, dass die Bundesnetzagentur jedes klinische System allein beaufsichtigt, noch dass jede Einrichtung dort eine verbindliche kostenlose Einzelfallfreigabe erhält.[26]

Für Medizinprodukte bietet das BfArM Informationen zu Qualifikation, Klassifizierung und Vorkommnismeldungen. Datenschutzfragen gehören zur jeweils zuständigen Datenschutzaufsicht. Eine Beratung durch das klinische Ethikkomitee ersetzt keinen dieser Wege.[18,26,35]

6. Wer verantwortlich sein sollte – und warum die Pflegeleitung nicht fehlen darf

Die AEM-Standards sehen Ethikberatung auch als organisationsethische Unterstützung an. Sie betonen ein geklärtes Mandat, multiprofessionelle Kompetenz, inhaltliche Unabhängigkeit und Ressourcen. Die AEM-Empfehlungen zur

Leitlinienentwicklung unterscheiden wiederum die Erarbeitung einer Empfehlung von ihrer autorisierten Inkraftsetzung.[14,15]

Daran anschließend schlagen wir folgende Aufgabenverteilung vor. Sie ist ein Organisationsmodell, keine allgemeingültige gesetzliche Zuständigkeitsordnung:

Rolle	Aufgabe im vorgeschlagenen KI-Verfahren	Grenze der Rolle
Träger, Geschäftsführung oder Vorstand	Zweck, Budget, Verantwortliche und Freigabebedingungen festlegen; verbleibende Organisationsrisiken begründet entscheiden.	Verantwortung nicht vollständig an Anwender oder Ethikstelle delegieren.
Pflegerische Leitung	Auswirkungen auf Pflegeprozess, Schichtbetrieb, Alarmbearbeitung, Personalressourcen und Beziehungsgestaltung prüfen; Umsetzung mitgestalten.	Nicht auf nachträgliche „Akzeptanzbeschaffung“ reduzieren.
Pflegefachpersonen und andere unmittelbar Anwendende	Alltagstauglichkeit, Fehler und unvorhergesehene Folgen in Tests und Rückmeldungen sichtbar machen.	Keine persönliche Sicherheitsgarantie für unbekannte Produktmängel übernehmen.
Ärztliche beziehungsweise sonstige fachliche Leitung	Fachliche Eignung, Vergleichsstandard und angemessene Reaktion auf Systemausgaben prüfen.	Nicht allein über organisatorische oder grundrechtliche Risiken entscheiden.
Ethikbeauftragte oder Ethikgremium	Werte- und Interessenkonflikte strukturieren, Betroffene beteiligen, Empfehlungen und Dissens dokumentieren.	Keine automatische Zulassungs-, Datenschutz- oder Haftungsinstanz.
Datenschutz, Informationssicherheit, Recht und gegebenenfalls Medizinproduktekompetenz	Jeweiligen Fachbereich prüfen und Anforderungen dokumentieren.	Getrennte Prüfungen ersetzen sich nicht; Datenschutzbeauftragte beraten und überwachen in ihrer gesetzlichen Rolle.
Beschäftigten- sowie Patienten- oder Nutzervertretung	Belastungen, verständliche Informationen, Beschwerdewege und Verteilungsfolgen einbringen.	Beteiligung nicht mit Zustimmung aller Betroffenen gleichsetzen.

Im Konzern treten an die Stelle der klinischen Rollen die fachlich verantwortlichen Geschäftsbereiche, etwa Personal, Kundenservice oder Produktion. Bei länderübergreifendem Einsatz sind lokale Perspektiven und Rechtslagen einzubeziehen. Ein zentrales Gremium sollte nicht verdecken, wer einen konkreten Prozess tatsächlich betreibt.

Die pflegerische Leitung ist besonders wichtig, weil die Wirkung einer Technik häufig erst zwischen Berufsgruppen und Schichten entsteht. Ein Alarm kann ärztlich sinnvoll und dennoch pflegerisch nicht bearbeitbar sein. Ein Roboter kann Wege sparen, aber zusätzliche Übergabe- und Aufsichtsarbeit erzeugen. Diese Konflikte sollten vor Beschaffung und Prozessfestlegung untersucht werden, nicht erst nach einer Produktdemonstration.

Die ZEKO betont bereits in ihrer Stellungnahme von 2021, dass die Verantwortung nicht allein auf einzelne Ärzte verlagert werden darf; auch Auswahl, Vorbereitung der Mitarbeitenden, Wartung und Schutz vor Missbrauch betreffen die Einrichtung.[13] Für die heutige Umsetzung empfehlen wir zusätzlich klar dokumentierte Eskalationsrechte, Zugang zu relevanten Unterlagen und geschützte Zeit für Prüfung und Schulung. Ein gesetzliches Universalveto der Ethikbeauftragten wird damit nicht behauptet.

Klinische Ethikberatung und Forschungsethik sind schließlich zu unterscheiden. Ob ein Vorhaben eine gesonderte Bewertung durch eine zuständige Forschungsethikkommission und weitere Genehmigungen benötigt, ist eigenständig zu klären. Die Bezeichnung „Pilot“ hebt entsprechende Anforderungen nicht auf.

7. Von Prinzipien zu Entscheidungen: ein Verfahren mit sechs Prüfstufen

Abbildung 3 – Sechs Prüfstufen vor und nach einer Freigabe

Stufe 1 · Auftrag, Vergleich und Grenzen festlegen Was soll besser werden, verglichen womit?
Stufe 2 · Belege statt Vorführungen verlangen Endpunkte, Population, Vergleichsgruppe
Stufe 3 · Rechte, Betroffene und Zielkonflikte prüfen Datenschutz, Beteiligung, Verteilungsfolgen
Stufe 4 · Begrenzten Pilot mit vorab festgelegten Kriterien planen Abbruchgründe vor dem Start
Stufe 5 · Begrenzte, begründete Freigabe erteilen Umfang, Dauer, Auflagen benennen
Stufe 6 · Betrieb, Änderungen und Ausstieg organisieren Rückkehr zum Alternativverfahren möglich halten

Gliederung nach Abschnitt 7 dieses Leitfadens. Redaktioneller Verfahrensvorschlag, kein wissenschaftlich validiertes Gesamtverfahren.

Die folgenden Stufen sind unsere redaktionelle Umsetzungsempfehlung. Sie greifen Gedanken aus Risikomanagement und früher klinischer Evaluation auf, sind aber selbst nicht als Gesamtverfahren validiert.[16,28] Prüftiefe und Aufwand sollen mit Schadenspotenzial, Reichweite, Autonomie und Unumkehrbarkeit wachsen. Ein interner Formulierungsvorschlag braucht nicht denselben Prozess wie eine Therapiepriorisierung.

Stufe 1: Auftrag, Vergleich und Grenzen festlegen

Die verantwortliche Stelle formuliert den Bedarf ohne Produktamen: Welches Problem soll für wen verbessert werden? Wie wird es heute gelöst? Welche nicht KI-basierten Alternativen bestehen? Anschließend werden zulässige Aufgabe, Nutzer, Daten und ausgeschlossene Verwendungen festgehalten. Ohne diese Beschreibung ist auch die Nutzenprüfung unscharf.

Die Organisation erstellt dafür ein Anwendungsregister. Es soll auch KI-Funktionen innerhalb bereits genutzter Standardsoftware erfassen. Ein risikoadäquater Meldeweg für informelle Nutzung ist hilfreicher als ein Register, das nur offiziell beschaffte Großprojekte kennt.

Stufe 2: Belege statt Vorführungen verlangen

Hersteller sollten Nachweise für die angebotene Version und den beabsichtigten Einsatz vorlegen. Zu prüfen sind Vergleichsgruppe, Datenherkunft, externe Validierung, relevante Endpunkte, Unsicherheiten, Finanzierung und Interessenkonflikte. Eine Veröffentlichung mit Autorenbezug zum Hersteller ist nicht automatisch unbrauchbar; die Abhängigkeit gehört aber in die Bewertung.

Je nach Frage kommen unterschiedliche Methoden infrage. Eine externe Validierung prüft Übertragbarkeit von Vorhersagen, eine geeignete vergleichende Wirkungsstudie den Nutzen des Einsatzes. Randomisierung ist für manche kausalen Fragen besonders hilfreich, aber nicht für jede Prüfaufgabe sinnvoll oder möglich. Methodische Beratung sollte deshalb vor Auswahl von Design und Fallzahl erfolgen.

Stufe 3: Rechte, Betroffene und Zielkonflikte prüfen

Hier laufen Rechts-, Datenschutz-, Sicherheits- und Ethikprüfung parallel. Erfasst werden auch indirekt Betroffene: Menschen, die weniger Aufmerksamkeit erhalten, Beschäftigte, deren Arbeit verdichtet wird, oder Gruppen, die vom digitalen Zugang ausgeschlossen bleiben.

Die Ethikberatung sollte nicht lediglich fragen, ob die Beteiligten „KI akzeptieren“. Sie sollte prüfen, welche Alternativen sie haben, ob sie widersprechen können und wer mögliche Nachteile trägt. Gesetzlich erforderliche Folgenabschätzungen werden identifiziert; darüber hinaus kann die Organisation eine freiwillige ethische Folgenprüfung vorsehen.

Stufe 4: Einen begrenzten Pilot mit vorab festgelegten Kriterien planen

Vor dem Einsatz werden Nutzenziel, Sicherheitsendpunkte, Messdauer, Datenzugriff und Abbruchkriterien festgelegt. Ein zunächst stiller Parallelbetrieb kann bei geeigneten Systemen prüfen, welche Ausgaben entstanden wären, ohne

unmittelbar Entscheidungen zu steuern. Er beweist allerdings noch nicht, wie Menschen auf die Ausgaben reagieren oder ob sich Ergebnisse verbessern.

In einem anschließenden kontrollierten Einsatz sind auch Korrekturarbeit, Verzögerungen, übersehene Warnungen und Auswirkungen auf andere Aufgaben zu erfassen. Es sollten nicht nur frühe Enthusiasten teilnehmen. Nachtdienst, Vertretung, Stress, Sprachbarrieren und technische Ausfälle gehören je nach Anwendung zum Testkontext.

Stufe 5: Eine begrenzte, begründete Freigabe erteilen

Eine Freigabe benennt Zweck, Version, Datenquellen, Nutzerkreis, Aufsicht, Ausschlüsse und nächstes Prüfdatum. Sie dokumentiert erwarteten Nutzen, verbleibende Risiken, offene Evidenzlücken und abweichende Voten. Eine Organisation kann eine Anwendung freigeben, weiter begrenzt erproben, zurückstellen oder ablehnen. Die Entscheidung und ihre Bedingungen sollten voneinander unterscheidbar sein.

Unsere Empfehlung ist, wesentliche Interessenkonflikte offenzulegen und bei besonders folgenreichen Einsätzen unabhängige Expertise einzubeziehen. Diese kann ein interner Bereich ohne Projektverantwortung oder eine externe Stelle sein. Eine einzige zusätzliche Unterschrift ist noch keine unabhängige Prüfung.

Stufe 6: Betrieb, Änderungen und Ausstieg organisieren

Nach der Einführung werden sowohl Leistung als auch Folgen überwacht. Auslöser einer Neubewertung sind beispielsweise Modell- oder Datenänderungen, ein anderer Nutzerkreis, neue Aufgaben, auffällige Fehler oder veränderte Arbeitsbedingungen. Nicht jedes Sicherheitsupdate verlangt einen vollständigen Neustart; die Tiefe der erneuten Prüfung ist nach Art und Tragweite der Änderung zu begründen.

Vorfälle brauchen eine erreichbare Meldestelle, klare Zuständigkeit und einen Weg zurück zur sicheren Versorgung oder Arbeitsweise. Für Beschwerden empfehlen wir einen benannten Empfänger, eine nach Dringlichkeit festgelegte Bearbeitungsfrist und eine fachlich befugte Überprüfung mit Korrekturmöglichkeit. Dringlicher Schadensverdacht braucht einen sofort nutzbaren Eskalationsweg. Eine Pause sollte für Betroffene nicht unbemerkt zum Versorgungsabbruch werden. Ausfall- und Ausstiegsverfahren sind deshalb praktisch zu testen. Hinweise von Beschäftigten dürfen nicht durch eine Kultur sanktioniert werden, die ausschließlich auf Einführungsziele und Nutzungsquoten schaut.

8. Was sinnvoll gemessen werden sollte

Warum „95 Prozent Genauigkeit“ wenig sagen kann

Ein bewusst vereinfachtes Rechenbeispiel, keine Studienzahl: Unter 10.000 untersuchten Personen haben 100 den gesuchten Zustand. Bei 90 Prozent Sensitivität werden 90 davon erkannt. Bei 95 Prozent Spezifität werden von den 9.900 übrigen Personen 495 fälschlich positiv markiert. Damit sind nur 90 von 585 positiven Meldungen richtig: rund 15,4 Prozent.

Das System kann trotzdem nützlich sein, wenn die anschließende Prüfung angemessen und der Nutzen ausreichend groß ist. Aber seine Folgen hängen entscheidend von Häufigkeit des Zustands, Fehlerkosten und verfügbarer Bearbeitungskapazität ab. Eine Kennzahl allein beantwortet diese Fragen nicht.

Dimension	Beispiele für geeignete Messgrößen	Typische Fehlinterpretation
Relevanter Nutzen	Patientenergebnis, korrekte Fallbearbeitung, nachvollziehbar bessere Dienstleistung	Modellscore wird mit Versorgungserfolg gleichgesetzt.
Sicherheit	Schweregrad von Fehlern, übersehene kritische Ereignisse, unnötige Folgeaktionen	Viele harmlose richtige Antworten überdecken seltene schwere Fehler.
Arbeitsablauf	Nettozeit einschließlich Kontrolle, Unterbrechungen, Reaktionszeit, Zusatzaufgaben anderer Teams	Eingesparte Schreibzeit wird als gesamte Arbeitersparnis ausgegeben.
Verteilung	Gruppenbezogene Ergebnisse, Zugangsbarrieren, ungleiche Belastung	Ein guter Durchschnitt gilt als Beleg für Fairness.

Dimension	Beispiele für geeignete Messgrößen	Typische Fehlinterpretation
Selbstbestimmung und Beziehung	Verständnis der Rolle des Systems, erreichbare Menschen, Beschwerden, berichtete Wahlmöglichkeiten	Hohe Nutzung wird mit freiwilliger Zustimmung verwechselt.
Robustheit und Betrieb	Ausfälle, relevante Veränderungen, Wiederherstellungszeit, Abhängigkeiten	Eine einmalige Abnahme gilt als dauerhafte Sicherheit.

Diese Tabelle enthält eigene Messvorschläge, keine universellen Mindestwerte. Schwellen müssen aus Schadensschwere, Baseline und Ziel des jeweiligen Vorhabens begründet werden. Beobachtungsdauer und Stichprobengröße brauchen methodische Planung. „Kein beobachteter Schaden“ in einem kleinen Pilot ist kein Beweis, dass seltene schwere Schäden ausgeschlossen sind.

Für sprachgenerierende Systeme sollten fachlich definierte Fehlertypen statt bloßer sprachlicher Gefälligkeit bewertet werden. Bei kritischen Inhalten sind nachvollziehbare Bewertungsregeln, geeignete unabhängige Beurteilungen und der Umgang mit Uneinigkeit zu planen. TRIPOD-LLM unterstützt die transparente Beschreibung solcher Untersuchungen, ersetzt aber keine Qualitätsprüfung.[29]

9. Drei illustrative Anwendungssituationen

Die folgenden Fälle sind konstruiert. Sie berichten keine tatsächlichen Einführungsprojekte und enthalten keine Produktfreigabe.

Krankenhaus: Ein Frühwarnsystem erzeugt mehr Hinweise als das Team bearbeiten kann

Ein Anbieter verspricht frühere Erkennung einer Verschlechterung. Vor der Einführung verlangt die Klinik nicht nur eine Leistungskennzahl, sondern eine Abschätzung der erwarteten Alarme pro Schicht, der erforderlichen Folgehandlungen und der übersehenen Ereignisse. Ärztliche und pflegerische Leitung planen gemeinsam, wer reagiert und welche Aufgaben dadurch zurückgestellt würden.

Die Ethikbeauftragte fragt insbesondere: Verbessert die Priorisierung tatsächlich die Versorgung, oder verschiebt sie knappe Aufmerksamkeit lediglich? Werden bestimmte Patientengruppen regelmäßig übersehen? Gibt es einen erreichbaren Eskalationsweg? Ein Pilot wird erst vertretbar, wenn Reaktionsfähigkeit und Messplan zur Aufgabe passen. Eine bessere Vorhersage ohne handlungsfähigen Prozess ist noch kein Versorgungskonzept.

Krankenhaus: Ein Assistent hört das Behandlungsgespräch mit

Das Ziel ist Entlastung bei der Dokumentation. Vorab werden die Datenflüsse geprüft: Was wird aufgezeichnet, wo verarbeitet, wer erhält Zugriff, wie lange wird gespeichert und wird Material für andere Zwecke verwendet? Ob eine Einwilligung erforderlich beziehungsweise eine andere Rechtsgrundlage tragfähig ist, wird fallbezogen geklärt. Eine beiläufige Zustimmung ist kein Ersatz für diese Prüfung. Unabhängig davon sind bei Tonaufnahmen die Vertraulichkeit des nichtöffentlich gesprochenen Wortes (§ 201 StGB) und bei externer Verarbeitung die berufliche Schweigepflicht einschließlich der Voraussetzungen für mitwirkende Dienstleister (§ 203 StGB) gesondert zu prüfen.[37]

Im Test interessiert nicht nur die Länge der erzeugten Notiz, sondern ob relevante Informationen fehlen oder hinzugefügt werden und wie viel Korrektur erforderlich ist. Als ethische Gestaltung empfehlen sich verständliche Information und eine praktikable Alternative, wenn Menschen die Aufzeichnung ablehnen. Ein allgemeines, ausnahmsloses gesetzliches Recht auf jede gewünschte analoge Alternative wird damit nicht behauptet.

Konzern: Ein System priorisiert Bewerbungen

Eine kürzere Bearbeitungszeit ist noch kein Nachweis besserer Auswahl. Das Unternehmen prüft, ob historische Auswahlentscheidungen vorhandene Benachteiligungen als Lernziel reproduzieren könnten und ob Kriterien für die ausgeschriebene Tätigkeit sachlich relevant sind. Sprachstil oder Lücken im Lebenslauf dürfen nicht unbesehen als Eignungsmaß dienen.

Zusätzlich ist die Hochrisikoeinstufung nach Art. 6 Abs. 2 in Verbindung mit Anhang III Nr. 4 Buchst. a einschließlich der begrenzten Ausnahme nach Art. 6 Abs. 3 zu prüfen. Dass am Ende ein Mensch entscheidet, beseitigt diese Frage

nicht; die einschlägigen Hochrisikovorschriften werden grundsätzlich ab 2. Dezember 2027 anwendbar.[20,21]

Die fachliche Verantwortung bleibt identifizierbar; Beschäftigtenvertretung, Recht, Datenschutz und Ethik werden ihrem Mandat entsprechend beteiligt. Betroffene benötigen die jeweils erforderlichen Informationen und wirksame Möglichkeiten, Fehler überprüfen zu lassen. Eine Funktion zur biometrisch gestützten Emotionsableitung im Bewerbungsinterview ist nicht einfach eine weitere Komfortfunktion: Hier ist insbesondere das Verbot nach Art. 5 zu prüfen. Eine menschliche Schlussunterschrift heilt weder eine verbotene Funktion noch automatisch eine diskriminierende Vorauswahl.[20,22,25]

10. Wo und wie Ethikbeauftragte sich informieren können

Die folgende Wegweisung soll Informationsarbeit verkürzen. Kein einzelnes Dokument deckt alle Fragen ab. Bei Internetseiten sind Fassungsdatum und Rechtsstand zu prüfen; ein offizieller Herausgeber garantiert nicht, dass jede Unterseite bereits an die neueste Gesetzesänderung angepasst wurde.

Konkrete Frage	Geeigneter Einstieg	Was damit zu tun ist – und was er nicht leistet
Wie sollte mein Ethikmandat ausgestaltet sein?	AEM: Standards und Empfehlungen [14,15]	Geschäftsordnung, Kompetenz, Unabhängigkeit und Ressourcen abgleichen; keine KI-Zulassung.
Was verlangt personenorientierte Pflege?	ICN-Ethikkodex, Volltext , besonders 1.11 [12]	Pflegeethische Ziele in lokale Anforderungen übersetzen; der Kodex ersetzt kein nationales Recht.
Wie verteilen sich klinische Verantwortungen?	ZEKO: KI-Entscheidungsunterstützung [13]	Rollen und fachliche Verantwortung diskutieren; Rechtsangaben aus 2021 gesondert aktualisieren.
Welche ethischen Aspekte werden leicht übersehen?	WHO: KI-Governance und generative Modelle [1,2]	Lokale Prüfung auf Autonomie, Gerechtigkeit und weitere Folgen kontrollieren; keine Einzelproduktbewertung.
Welche KI-Pflichten gelten wann?	AI Act bei EUR-Lex und Änderungsverordnung 2026/1744 [20,21]	Rechtsakt, Änderungen, Geltungsbereich und Anwendungsdatum zusammen lesen.
Welche deutsche Stelle ist zuständig?	KI-MIG § 2 und zentrale Anlaufstelle § 6 [26]	Zuständigkeit anhand System und Betreiber klären; nicht jede Anfrage gehört allein an BNetzA oder BfArM.
Wie prüfe ich personenbezogene Datenverarbeitung?	DSK: Orientierungshilfen [18,19]	Datenflüsse und Prüffragen mit Datenschutzbeauftragten durcharbeiten; konkrete Aufsicht je nach Träger klären.
Handelt es sich um ein Medizinprodukt?	BfArM: rechtlicher Status und Klassifizierung , MDR und MDCG [23,35,36]	Zweckbestimmung und Einstufung fachkundig prüfen; nicht nur auf die Vertriebsbezeichnung vertrauen.
Wie werden medizinproduktebezogene Risiken gemeldet?	BfArM: Risiken melden [35]	Den einschlägigen Meldeweg in die internen Vorfallsprozesse integrieren.
Wie strukturiere ich KI-Risikomanagement?	NIST AI RMF und GenAI-Profil [16,17]	Aufgaben zuordnen und Nachweise planen; freiwillige Rahmenwerke sind kein EU-Konformitätsnachweis.
Welche technischen Angriffsmöglichkeiten sind zu prüfen?	BSI: Künstliche Intelligenz [34]	IT-Sicherheit mit konkreter Architektur und Rechten befassen; kein allgemeiner Prompt ersetzt Sicherheitstests.
Wie prüfe ich die Qualität einer Studie?	DECIDE-AI, TRIPOD-LLM, STANDING Together [28–30]	Bericht und Datengrundlage gezielt prüfen; Checklisten sind keine Gütesiegel. Biometrische und fachliche Beratung hinzuziehen.
Was bedeutet ein KI-Managementzertifikat?	ISO/IEC 42001: Erläuterung [31]	Geltungsbereich und Ausschlüsse des Zertifikats prüfen; Produktnutzen zusätzlich belegen lassen.

Für eigene Literaturrecherchen ist PubMed ein sinnvoller Einstieg in medizinische Studien. Eine Suchanfrage sollte nicht nur „AI ethics“ enthalten, sondern konkrete Aufgabe, Produkt oder Modell, Population und Evaluationstyp. Beispielhaft: ("artificial intelligence" OR "machine learning") AND sepsis AND ("external validation" OR randomized). Produktnamen und aktuelle Versionen sind ergänzend zu suchen.

Anschließend sind Originalarbeit, Studienregistrierung soweit vorhanden, Finanzierung, Korrekturen und spätere Untersuchungen zu prüfen. Abstracts helfen beim Auffinden, reichen aber für anspruchsvolle Beschaffungsentscheidungen häufig nicht aus. Wenn die eigene Einrichtung diese Beurteilung nicht leisten kann, ist methodische Unterstützung durch Biometrie, klinische Epidemiologie oder fachlich geeignete externe Einrichtungen sinnvoll. Das ist kein Kompetenzmangel der Ethikstelle, sondern Teil einer guten Aufgabenverteilung.

11. Was Gesellschaft und Organisationen als Lösung anstreben sollten

Die folgenden Punkte sind normative Schlussfolgerungen dieses Beitrags, keine Ergebnisse einer einzelnen Studie.

Erstens sollten Nutzen und Risiken gemeinsam bewertet werden. Wer nur Gefahren beschreibt, unterschätzt mögliche Verbesserungen. Wer nur Effizienz zählt, übersieht Rechte, Beziehungen und ungleich verteilte Belastungen. Beides kann zu schlechten Entscheidungen führen.

Zweitens sollte mit Autonomie und Reichweite die Nachweispflicht wachsen. Ein System, das Entwürfe erstellt, benötigt andere Begrenzungen als eines, das unmittelbar über Ressourcen verfügt oder irreversible Handlungen ausführt. Es ist weder nötig, jede harmlose Automatisierung manuell abzunicken, noch vertretbar, wirksame Kontrolle nur auf dem Papier vorzusehen.

Drittens sollten Betroffene Mitsprache und Abhilfe erhalten. Dazu gehören verständliche Informationen, erreichbare Menschen und eine Überprüfung folgenreicher Fehler. Institutionelle Verantwortung darf nicht in einer langen Kette aus Hersteller, Integrator und Anwender verschwinden.

Viertens sollten Gewinne bewusst verteilt werden. Freigesetzte Zeit kann bessere Betreuung, Weiterbildung oder geringere Überlastung ermöglichen. Ob sie tatsächlich dafür genutzt wird, ist eine Leitungs- und Beteiligungsentscheidung. Ein Produktivitätsgewinn beantwortet diese Verteilungsfrage nicht von selbst.

Fünftens braucht es öffentliche und unabhängige Kompetenz: Forschung über tatsächliche Wirkungen, verlässliche Aufsicht, fachlich belastbare Prüfkapazität und Austausch über Vorfälle. Nicht jede einzelne Klinik und nicht jedes Unternehmen sollte dieselben grundlegenden Sicherheitsfragen isoliert neu lösen müssen.

Die tragfähige Lösung ist deshalb kein einzelnes Ethiksigel und keine perfekte Maschine. Es ist eine lernfähige Verantwortungsordnung: konkrete Vorteile nachweisen, Grenzen offenlegen, Menschen wirksam beteiligen und Entscheidungen korrigieren können. KI sollte menschliche Handlungsmöglichkeiten erweitern. Ob sie das tut, muss sich an den Erfahrungen und Ergebnissen der betroffenen Menschen zeigen – nicht an der Überzeugungskraft einer Demonstration.

Die folgenden Vorlagen sind eigene Organisationsvorschläge. Sie werden erst durch die jeweils zuständige Einrichtung und im Rahmen des geltenden Rechts zu internen Vorgaben.

A. Die ersten 90 Tage: eine realistische Arbeitsstruktur aufbauen

Der Zeitraum ist ein Planungsbeispiel, keine gesetzliche Frist und kein Versprechen, innerhalb von 90 Tagen jede Anwendung freigeben zu können.

Zeitraum	Priorität	Greifbares Ergebnis
Tag 1–15	Mandat, Ansprechpartner und bereits verwendete Systeme erfassen.	Vorläufiges Register; verantwortliche Leitung; erreichbarer Meldeweg.
Tag 16–30	Anwendungen nach Folgen, Daten, Autonomie und Reichweite priorisieren.	Begründete Prüfreihefolge; sofortiger Umgang mit dringlichen Sicherheits- oder Rechtsproblemen.

Zeitraum	Priorität	Greifbares Ergebnis
Tage 31–60	Für priorisierte Anwendungen Evidenz, Rechte, Arbeitsabläufe und Betroffenenperspektiven prüfen.	Entscheidungsunterlagen; gegebenenfalls methodisch geplanter, zulässiger Pilot.
Tage 61–90	Erste Entscheidungen dokumentieren; Vorfalls- und Änderungsprozesse erproben.	Begrenzte Freigaben oder begründete Zurückstellungen; Prüfkalender; Ausfallübung.

Ein laufender Einsatz mit ernstem Schadensverdacht wartet nicht bis Tag 61. Er benötigt unverzügliche Bewertung und gegebenenfalls eine sichere Aussetzung mit Ersatzverfahren. Umgekehrt darf der Zeitplan nicht dazu führen, dass unzureichende Daten aus einem kurzen Pilot als ausreichender Sicherheitsnachweis erklärt werden.

B. Zwölf Fragen, die vor einer Beschaffung beantwortet sein sollten

1. Welchen genau abgegrenzten Zweck decken Produkt und Nachweise ab, und welche Verwendung ist ausdrücklich ausgeschlossen?
2. Welche Modell-, Software- und Datenversion wurde untersucht, und entspricht sie der angebotenen Konfiguration?
3. Welche unabhängige oder externe Evaluation liegt vor; wie wurden Vergleichsgruppe und relevante Endpunkte gewählt?
4. Welche bekannten Fehler, gefährlichen Auslassungen und Grenzen bestehen, auch bei seltenen Situationen?
5. Wie variiert die Leistung zwischen relevanten Gruppen und Standorten; wo fehlen belastbare Daten?
6. Welche personenbezogenen und vertraulichen Informationen werden verarbeitet, weitergegeben, gespeichert oder für zusätzliche Zwecke verwendet?
7. Welche Handlungen kann das System selbst ausführen, und wie werden Zugriffe, fremde Inhalte und Manipulationsversuche begrenzt?
8. Welche konkrete Arbeit verbleibt bei Nutzern, wer bearbeitet Warnungen, und welcher Schulungs- und Kontrollaufwand entsteht?
9. Welche Protokolle, Auswertungen und Informationen über Fehler erhalten Betreiber tatsächlich?
10. Wie werden Modell- und Produktänderungen angekündigt, bewertet und gegebenenfalls zurückgenommen?
11. Welche vertraglichen Regelungen bestehen zu Unterstützung, Vorfällen, Prüfzugang, Verfügbarkeit, Datenexport und Vertragsende?
12. Welche sichere Alternative steht bei Ausfall, unvertretbarer Leistung oder Anbieterwechsel zur Verfügung?

Fehlende Antworten sind nicht automatisch ein endgültiges Verbot. Sie bestimmen aber, welche Unsicherheit noch besteht und ob sie für den konkreten Zweck vertretbar ist. Verbotene Verwendungen oder fehlende zwingende Voraussetzungen lassen sich nicht durch einen hohen erwarteten wirtschaftlichen Nutzen ausgleichen.

C. Inhalt einer nachvollziehbaren Entscheidungsvorlage

Eine angemessen kurze Vorlage sollte folgende Felder enthalten:

- **Anwendung:** Bezeichnung, konkrete Aufgabe, Standort, Version, Nutzerkreis, Verantwortliche; rechtliche Rolle der Einrichtung sowie relevante Eigenentwicklungen oder Veränderungen.
- **Vergleich und Ziel:** Ausgangslage, nichttechnische Alternativen, erwarteter menschlicher Nutzen.
- **Betroffene:** direkt und indirekt betroffene Gruppen, Beteiligung, wesentliche Einwände.
- **Evidenz:** wichtigste Nachweise, lokale Übertragbarkeit, ungeklärte Fragen, Interessenkonflikte.
- **Recht und Daten:** einschlägige Einordnung, geprüfte Grundlagen, erforderliche Beteiligungen und Folgenabschätzungen.
- **Sicherheit und Betrieb:** Fehlerfolgen, Rechte, Aufsicht, Kapazitäten, Ausfall- und Ersatzverfahren.
- **Entscheidung:** freigegeben, begrenzt erproben, zurückgestellt oder abgelehnt; jeweils mit Begründung.
- **Bedingungen:** erlaubte Verwendung, Ausschlüsse, Messgrößen, Schwellen, Zuständigkeit für Eingriffe.

- **Fortschreibung:** nächster Prüftermin und anlassbezogene Auslöser.
- **Dokumentation:** Entscheidungsträger, abweichende Voten, Datum und zugängliche Fassung.

Ein Freigabedokument sollte beispielsweise nicht nur „Ethik geprüft“ festhalten, sondern den tatsächlichen Konflikt und seine Bearbeitung benennen: etwa, warum eine zusätzliche Warnbelastung angesichts eines erwarteten Nutzens vertretbar ist und wie sie begrenzt werden soll.

D. Wann eine Anwendung nicht unverändert weiterlaufen sollte

Beobachtung	Empfohlenes Vorgehen
Plausibler Verdacht auf schweren Schaden oder eine verbotene Funktion	Unverzögliche qualifizierte Bewertung, gegebenenfalls sichere Aussetzung; einschlägige Meldepflichten prüfen und erfüllen.
Datenzugriff oder Weitergabe außerhalb des vorgesehenen Rahmens	Sicherheits- und Datenschutzprozess aktivieren; Zugriff begrenzen, Sachverhalt sichern und bewerten.
Relevante Leistungsverschlechterung oder neue Benachteiligung	Ursachenanalyse und begründete Begrenzung, Anpassung oder Pause nach Schadenspotenzial.
Warnungen bleiben mangels Personal unbearbeitet	Arbeitsprozess und Kapazität ändern; Nutzenannahme und Freigabe neu bewerten.
Modell, Zweck oder Nutzerkreis ändert sich wesentlich	Änderung nicht stillschweigend von der alten Freigabe abdecken lassen.
Keine sichere Versorgung oder Arbeit bei Ausfall möglich	Abhängigkeit reduzieren und tragfähigen Ersatz vorbereiten, bevor weitere Ausweitung erfolgt.

Die konkrete Reaktion sollte verhältnismäßig sein und Risiken eines abrupten Abschaltens einbeziehen. Gesetzliche Melde- und Handlungspflichten gehen vor; die Tabelle behauptet keine einheitliche Rechtsfolge für alle Fälle.

E. Kleines Glossar

Sensitivität: Anteil tatsächlich positiver Fälle, die das System erkennt. **Spezifität:** Anteil tatsächlich negativer Fälle, die es korrekt nicht als positiv markiert. **Positiver Vorhersagewert:** Anteil tatsächlich positiver Fälle unter allen positiven Meldungen; stark von der Häufigkeit des Zustands abhängig. **AUROC:** Maß für die Rangtrennung positiver und negativer Fälle über verschiedene Schwellen; kein direkter Nachweis klinischen Nutzens. **Kalibrierung:** Übereinstimmung vorhergesagter Wahrscheinlichkeiten mit beobachteten Häufigkeiten. **Externe Validierung:** Prüfung an Daten außerhalb der Entwicklungsgrundlage. **Automation Bias:** unangemessene Orientierung an automatisierten Hinweisen. **RAG:** Abruf von Dokumenten oder Informationen zur Ergänzung des Modellkontexts; keine Wahrheitsgarantie. **DSFA:** Datenschutz-Folgenabschätzung. **FRIA:** Grundrechte-Folgenabschätzung nach dem einschlägigen KI-Rechtsrahmen. **Betreiber:** eine rechtliche Rolle, deren konkrete Zuordnung zu prüfen ist; nicht automatisch die Person, die eine ethische Stellungnahme verfasst.

Methodik, Grenzen und Hinweise vor Veröffentlichung

Rechercheansatz

Dieser Beitrag beruht auf einer gezielten narrativen Recherche bis zum 24. September 2026. Verwendet wurden Originalstudien, Veröffentlichungen der verantwortlichen Fachgesellschaften und Institutionen, amtliche Rechtsquellen und behördliche Orientierungshilfen. Gesucht wurde entlang der Themen klinischer Nutzen, Mensch-KI-Zusammenarbeit, Fehlalarme, Diskriminierung, Arbeitsproduktivität, Datenschutz, institutionelle Ethik und aktuelle KI-Regulierung. Vorrang hatten Quellen, die eine konkrete Behauptung tragen und ihre Grenzen erkennen lassen.

Die Auswahl umfasst positive, negative und gemischte Befunde. Es gab jedoch kein vorregistriertes Suchprotokoll, kein vollständiges Datenbankscreening und keine formale Verzerrungsbewertung aller Studien. Eine PRISMA-konforme systematische Übersicht oder Metaanalyse wird ausdrücklich nicht beansprucht. Auch wurden keine eigenen empirischen Daten erhoben. Insbesondere zu langfristiger Dequalifizierung, Beziehungseffekten,

Beschäftigungsfolgen und fortgeschrittenen autonomen Systemen ist damit keine umfassende Evidenzbewertung geleistet.

Bei teilweise eingeschränktem Volltextzugang wurden öffentlich zugängliche Originalabstracts und institutionelle Publikationsnachweise verwendet; die entsprechenden Nachweise sind unten gekennzeichnet. Bei technischen Abrufproblemen amtlicher Rechtsseiten wurden zusätzlich indexierte Passagen derselben amtlichen Quelle und offizielle Mitteilungen herangezogen. Eine nicht vollständig zugängliche Publikation wurde nicht als vollständig analysierter Volltext behandelt.

Bewusst getrennte Aussageebenen

Ebene	Beispiel im Beitrag	Was daraus nicht folgt
Empirischer Befund	Gemessene Zeitveränderung in einer bestimmten randomisierten Studie.	Gleicher Effekt bei allen Produkten und Einrichtungen.
Fachlicher oder ethischer Standard	ICN-Vorgabe zur personenzentrierten Pflege.	Automatisch unmittelbar geltendes deutsches Gesetz.
Rechtsnorm	Voraussetzungen einer DSFA nach Art. 35 DSGVO.	DSFA-Pflicht allein aufgrund der Bezeichnung „KI“.
Eigene Empfehlung	Sechsstufiges Entscheidungsverfahren und 90-Tage-Einstieg.	Validierte Leitlinie oder allgemeine gesetzliche Pflicht.

Redaktioneller Prüfstand

Zentrale Zahlen wurden gegen die benannten Originalnachweise abgeglichen. Besonders berücksichtigt wurden die Unterscheidung von Nichtunterlegenheit und Überlegenheit bei MASAI, von Modellvalidierung und klinischem Nutzen, von Dokumentationszeit und Gesamtarbeitszeit sowie ältere und aktuelle Produkt- beziehungsweise Publikationsfassungen. Rechtsakt und Anwendungsdatum werden auseinandergehalten. Das Zahlenbeispiel in Abschnitt 8 ist ausdrücklich selbst konstruiert.

Diese Fassung wurde mit KI-Unterstützung recherchiert und formuliert. Sie ist nicht extern fachlich begutachtet. Vor Veröffentlichung unter redaktioneller Autorenschaft empfehlen sich eine unabhängige klinisch-methodische und eine juristische Schlussprüfung, die Prüfung sämtlicher Links sowie die Offenlegung von Autorenschaft und relevanten Interessenkonflikten. Bei späterer Veröffentlichung sind insbesondere Rechtsstand und jüngere Wirksamkeitsstudien erneut zu prüfen. Es wird keine Beratung oder Prüfung durch eine reale Fachperson behauptet, die nicht stattgefunden hat.

Literatur und Primärquellen

Die Nummern beziehen sich auf die Belege im Text. Webquellen zuletzt für diese Fassung recherchiert am 24.09.2026. DOI-Links führen zur jeweiligen Originalpublikation. Bei Rechtsquellen sind konsolidierte Fassung und Änderungsrechtsakte gemeinsam zu beachten.

Empirische Forschung

[3] Gommers J et al. *Interval cancer, sensitivity, and specificity comparing AI-supported mammography screening with standard double reading without AI in the MASAI study*. The Lancet. 2026;407:505–514. DOI: [10.1016/S0140-6736\(25\)02464-X](https://doi.org/10.1016/S0140-6736(25)02464-X). Zugänglicher Originalabstract, Universität Lund. Für diesen Beitrag insbesondere der Originalabstract ausgewertet.

[4] Goh E et al. *Large Language Model Influence on Diagnostic Reasoning: A Randomized Clinical Trial*. JAMA Network Open. 2024;7:e2440969. DOI: [10.1001/jamanetworkopen.2024.40969](https://doi.org/10.1001/jamanetworkopen.2024.40969). Originalvolltext.

[5] Wong A et al. *External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients*. JAMA Internal Medicine. 2021;181:1065–1070. DOI: [10.1001/jamainternmed.2021.2626](https://doi.org/10.1001/jamainternmed.2021.2626). Originalvolltext. Historische Produktgeneration.

- [6] Wong A, Currey D, Schwinne M et al. *Multicenter Prospective Validation of an Updated Proprietary Sepsis Prediction Model*. JAMA Network Open. 2026;9(2):e260181. DOI: [10.1001/jamanetworkopen.2026.0181](https://doi.org/10.1001/jamanetworkopen.2026.0181). [Originalvolltext](#).
- [7] Lukac PJ et al. *Ambient AI Scribes in Clinical Practice: A Randomized Trial*. NEJM AI. 2025;2(12). DOI: [10.1056/AIoa2501000](https://doi.org/10.1056/AIoa2501000). PubMed; PMC. Originalabstract und öffentlich indexierte Originalpassagen ausgewertet, direkter Volltextabruf teilweise eingeschränkt.
- [8] Dratsch T et al. *Automation Bias in Mammography: The Impact of Artificial Intelligence BI-RADS Suggestions on Reader Performance*. Radiology. 2023;307:e222176. DOI: [10.1148/radiol.222176](https://doi.org/10.1148/radiol.222176). [Originalabstract im Hochschulrepositorium](#).
- [9] Obermeyer Z, Powers B, Vogeli C, Mullainathan S. *Dissecting racial bias in an algorithm used to manage the health of populations*. Science. 2019;366:447–453. DOI: [10.1126/science.aax2342](https://doi.org/10.1126/science.aax2342). [Originalartikel als PDF bei der FTC](#).
- [10] Brynjolfsson E, Li D, Raymond LR. *Generative AI at Work*. Quarterly Journal of Economics. 2025;140:889–942. DOI: [10.1093/qje/qjae044](https://doi.org/10.1093/qje/qjae044). [Finale Originalfassung](#). Nicht mit den abweichenden Zahlen früher Arbeitspapiere vermischen.
- [11] Dell’Acqua F et al. *Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of Artificial Intelligence on Knowledge Worker Productivity and Quality*. Organization Science. 2026;37:403–423. DOI: [10.1287/orsc.2025.21838](https://doi.org/10.1287/orsc.2025.21838). Finale Veröffentlichung vom 11.03.2026, nicht das frühere Arbeitspapier.

Ethik, Organisation, Sicherheit und Evaluationsmethodik

- [1] World Health Organization. *Ethics and governance of artificial intelligence for health*. 2021. [Originalpublikation](#); [offizielle Erläuterung der sechs Prinzipien](#).
- [2] World Health Organization. *Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models*. 2024. [Publikation](#); [Einführung vom 18.01.2024](#). Das abweichende Datum der aktuellen Landingpage ist kein Nachweis einer neuen Ausgabe.
- [12] International Council of Nurses. *The ICN Code of Ethics for Nurses*. 2021, insbesondere Abschnitt 1.11. [Vollständiger englischer Kodex](#).
- [13] Zentrale Ethikkommission bei der Bundesärztekammer. *Entscheidungsunterstützung ärztlicher Tätigkeit durch Künstliche Intelligenz*. 2021. DOI: [10.3238/arztebl.zeko_sn_cdss_2021](https://doi.org/10.3238/arztebl.zeko_sn_cdss_2021). [Stellungnahme und Original-PDF](#). Ethische Orientierung; nicht als aktueller AI-Act-Kommentar verwenden.
- [14] Vorstand der Akademie für Ethik in der Medizin. *Standards für Ethikberatung im Gesundheitswesen*. Verabschiedet am 21.12.2022, publiziert 2023. [Original-PDF](#).
- [15] Vorstand der Akademie für Ethik in der Medizin. *Empfehlungen zur Erstellung von Ethik-Leitlinien in Einrichtungen des Gesundheitswesens*. 2024. DOI: [10.1007/s00481-024-00820-x](https://doi.org/10.1007/s00481-024-00820-x). [Original-PDF](#).
- [16] National Institute of Standards and Technology. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. 2023. DOI: [10.6028/NIST.AI.100-1](https://doi.org/10.6028/NIST.AI.100-1). [Original-PDF](#). Freiwilliges Rahmenwerk; laufende Überarbeitung nicht mit veröffentlichter Neufassung verwechseln.
- [17] NIST. *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile*. 2024. DOI: [10.6028/NIST.AI.600-1](https://doi.org/10.6028/NIST.AI.600-1). [Original-PDF](#).
- [18] Datenschutzkonferenz. *Orientierungshilfe Künstliche Intelligenz und Datenschutz*. 06.05.2024. [Original-PDF](#). Behördliche Orientierung, nicht Gesetz.
- [19] Datenschutzkonferenz. *Orientierungshilfe zu datenschutzrechtlichen Aspekten von Retrieval-Augmented Generation*. Oktober 2025, Version 1.0. [Original-PDF](#).
- [28] Vasey B et al. *Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI*. Nature Medicine. 2022;28:924–933. DOI: [10.1038/s41591-022-01772-9](https://doi.org/10.1038/s41591-022-01772-9). Berichtsempfehlung; hier öffentlich zugängliche Verlagsinformationen ausgewertet.

- [29] Gallifant J et al. *The TRIPOD-LLM reporting guideline for studies using large language models*. Nature Medicine. 2025;31:60–69. DOI: [10.1038/s41591-024-03425-5](https://doi.org/10.1038/s41591-024-03425-5). Publikationsjahr 2025 trotz „2024“ im DOI.
- [30] Alderman JE et al. *Tackling algorithmic bias and promoting transparency in health datasets: the STANDING Together consensus recommendations*. The Lancet Digital Health. 2025;7(1):e64–e88. DOI: [10.1016/S2589-7500\(24\)00224-3](https://doi.org/10.1016/S2589-7500(24)00224-3). [Originalabstract](#). Konsensarbeit, keine Wirksamkeitsstudie; online erstmals am 18.12.2024 veröffentlicht.
- [31] ISO/IEC 42001:2023. *Artificial intelligence — Management system*. [Offizielle Standardbeschreibung](#); [offizielle Erläuterung](#). Hier wurden die frei zugänglichen Beschreibungen, nicht der kostenpflichtige Normvolltext ausgewertet.
- [32] *International AI Safety Report 2026*. Veröffentlichung 03.02.2026. [Vollständiger Bericht](#). Internationale wissenschaftliche Evidenzsynthese zu allgemeiner KI; keine verbindliche sektorale Richtlinie.
- [33] International Energy Agency. *Key Questions on Energy and AI*. 2026. [Executive Summary](#). Beobachtungen und Prognosen auseinanderhalten.
- [34] Bundesamt für Sicherheit in der Informationstechnik. [Themenbereich Künstliche Intelligenz](#); [Warnung zu indirekten Prompt-Injection-Angriffen](#). Ergänztender fachbehördlicher Einstieg; einzelne Direktabrufe eingeschränkt.

Recht und regulatorische Orientierung

- [20] Verordnung (EU) 2024/1689, AI Act, in der am Recherchetag maßgeblichen Fassung. [Ursprünglicher Rechtsakt](#). Insbesondere Art. 3–6, 25–27, 50, 111 und 113 sowie Anhänge I und III; zusammen mit [21] lesen.
- [21] Verordnung (EU) 2026/1744 vom 08.07.2026, Amtsblatt vom 24.07.2026, Digital Omnibus on AI. [Amtlicher Rechtsakt](#); [amtlicher englischer Text](#); [Kommission: Inkrafttreten am 27.07.2026](#). Die Änderung des Art. 113 verschiebt bestimmte Hochrisikovorschriften, nicht den gesamten AI Act.
- [22] Verordnung (EU) 2016/679, Datenschutz-Grundverordnung. [Amtlicher Rechtsakt](#). Insbesondere Art. 5, 6, 9, 13–15, 22, 28, 32 und 35; weitere Vorschriften können je nach Datenfluss relevant sein.
- [23] Verordnung (EU) 2017/745 über Medizinprodukte, MDR. [Amtlicher Rechtsakt](#). Insbesondere Art. 2, Erwägungsgrund 19 und Anhang VIII Regel 11; jeweils geltende Fassung prüfen.
- [24] Betriebsverfassungsgesetz. [§ 87](#), [§ 90](#), [§ 118](#). Bei öffentlichen und kirchlichen Trägern einschlägiges Vertretungsrecht gesondert klären.
- [25] Allgemeines Gleichbehandlungsgesetz. [Amtlicher Gesetzestext](#).
- [26] Gesetz zur Marktüberwachung und Innovationsförderung von künstlicher Intelligenz (KI-Marktüberwachungs-und-Innovationsförderungs-Gesetz – KI-MIG). In Kraft seit 29.07.2026. [Amtlicher Gesetzestext](#); [§ 2 Zuständigkeiten](#), [§ 6 Anlaufstelle](#), [§ 8 Beschwerdestelle](#), [§ 12 Information und Beratung](#).
- [27] Bürgerliches Gesetzbuch. [§ 630a Abs. 2](#), allgemein anerkannter fachlicher Standard bei medizinischer Behandlung.
- [35] Bundesinstitut für Arzneimittel und Medizinprodukte. [Rechtlicher Status und Klassifizierung](#); [Risiken melden](#). Fachbehördliche Einstiegs- und Meldeinformationen, keine allgemeine Freigabe sämtlicher Klinik-KI.
- [36] Medical Device Coordination Group. *MDCG 2019-11 rev.1: Qualification and classification of software*, Juni 2025. [Offizielle Veröffentlichung](#). Ergänzend *MDCG 2025-6*, Zusammenspiel von MDR/IVDR und AI Act: [offizielle Veröffentlichung](#). Unverbindliche fachbehördliche Orientierung; ältere Fristenangaben mit [21] abgleichen.
- [37] Strafgesetzbuch. [§ 201: Vertraulichkeit des Wortes](#); [§ 203: Verletzung von Privatgeheimnissen](#). Aufnahme, Verarbeitung und Einbindung externer Dienstleister anhand der konkreten Architektur und Berechtigung prüfen.

12 Rechtliche Hinweise

12.1 Keine Beratung

Diese Veröffentlichung ist ein redaktionelles Erzeugnis. Sie stellt keine Rechtsberatung, keine medizinische, pflegerische, arbeitsrechtliche oder datenschutzrechtliche Beratung und keine Beschaffungs- oder Vergabeberatung dar. Sie ersetzt insbesondere nicht die Prüfung durch Rechtsanwaltschaft, Datenschutzaufsicht, Medizinprodukteberatung, Vergabestelle, Ethikkommission nach Medizinprodukterecht oder Arbeitnehmervertretung. Wer auf Grundlage dieses Dokuments entscheidet, entscheidet auf eigene Verantwortung und hat den Sachverhalt vorher eigenständig fachlich prüfen zu lassen.

12.2 Keine Zusicherung, keine Konformitätsaussage

Die beschriebenen Verfahren, Prüfstufen und Checklisten sind redaktionelle Vorschläge. Sie sind keine Zusicherung einer Eigenschaft, kein garantiertes Ergebnis und kein Nachweis der Einhaltung einer Rechtsvorschrift. Wer ein Verfahren aus diesem Dokument anwendet, erfüllt damit keine gesetzliche Anforderung; ob eine Anforderung erfüllt ist, entscheidet die zuständige Stelle am Einzelfall.

12.3 Rechtsstand

Rechtsstände, Fristen und Fundstellen beziehen sich auf den im Dokumentkopf genannten Tag. Der KI-Rechtsrahmen wird laufend geändert; ein älterer Fristenplan darf nicht unverändert weiterverwendet werden. Genannte Gesetze und Verordnungen sind in ihrer jeweils geltenden Fassung heranzuziehen.

12.4 Haftung

Die Inhalte wurden mit Sorgfalt erstellt und, soweit im Methodenteil beschrieben, gegengeprüft. Für Richtigkeit, Vollständigkeit und Aktualität wird gleichwohl keine Gewähr übernommen. Eine Haftung des Herausgebers ist auf Vorsatz und grobe Fahrlässigkeit beschränkt; dies gilt nicht bei der Verletzung des Lebens, des Körpers oder der Gesundheit sowie bei zwingender gesetzlicher Haftung.

12.5 Verweise auf fremde Inhalte

Dieses Dokument verweist auf Inhalte Dritter, insbesondere auf Behörden-, Fachgesellschafts-, Normen- und Fachverlagsseiten. Auf deren Gestaltung und Inhalt hat der Herausgeber keinen Einfluss; für sie ist jeweils der dortige Anbieter verantwortlich. Die Adressen waren zum im Literaturverzeichnis genannten Prüfzeitpunkt erreichbar. Die Nennung eines Produkts, Herstellers oder Dienstes ist weder eine Empfehlung noch eine Bewertung seiner Eignung.

12.6 Marken und Produktbezeichnungen

Genannte Produkt-, Marken- und Unternehmensbezeichnungen sind Kennzeichen ihrer jeweiligen Inhaber und werden ausschließlich zur Beschreibung des Sachverhalts verwendet. Aus der Nennung folgt keine Verbindung zwischen dem Herausgeber und dem jeweiligen Unternehmen und keine Aussage über Zulassungs-, Konformitäts- oder Zertifizierungsstatus.

12.7 Urheberrecht und Nutzung

Dieses Dokument, sein Text, seine Tabellen und seine Abbildungen sind urheberrechtlich geschützt. Rechteinhaberin ist die WebTrades GmbH. Zulässig sind das Lesen, das Speichern und der Ausdruck für den eigenen Gebrauch sowie das Zitieren einzelner Stellen mit vollständiger Quellenangabe nach der im Impressum genannten Zitierweise. Die Vervielfältigung, Bearbeitung, Verbreitung, öffentliche Zugänglichmachung oder kommerzielle Verwertung des Dokuments oder wesentlicher Teile daraus bedarf der vorherigen schriftlichen Zustimmung der Rechteinhaberin. Das gilt ausdrücklich auch für die Übernahme von Tabellen und Abbildungen.

12.8 Keine personenbezogenen Daten

Für dieses Dokument wurden keine personenbezogenen Daten verarbeitet. Es beruht auf veröffentlichter Literatur, amtlichen Rechtsquellen und Veröffentlichungen von Fachgesellschaften; es wurden weder Beschäftigte befragt noch Patientendaten verwendet. Wer die hier beschriebenen Verfahren auf eine reale Einrichtung anwendet, verlässt diesen Rahmen: Die Auswertung von Tätigkeits-, Alarm- oder Quittierungsdaten unterliegt der Datenschutz-Grundverordnung und in der Regel der Mitbestimmung.

12.9 Unabhängigkeit

Dieses Dokument ist ohne Auftrag, ohne Finanzierung und ohne inhaltliche Beteiligung eines Herstellers, Krankenhausträgers, Kostenträgers, Verbands oder Beratungsunternehmens entstanden. Es besteht kein Vergütungs- oder Beteiligungsverhältnis zu einem der genannten Unternehmen oder Institutionen. Genannt werden Systeme, Anbieter und Dokumente, weil sie in den beschriebenen Gegenstand fallen, nicht weil sie beworben werden sollen.

12.10 Anwendbares Recht

Es gilt das Recht der Bundesrepublik Deutschland unter Ausschluss des UN-Kaufrechts. Gesetzliche Verbraucherschutzvorschriften bleiben unberührt.