Universität Augsburg
University

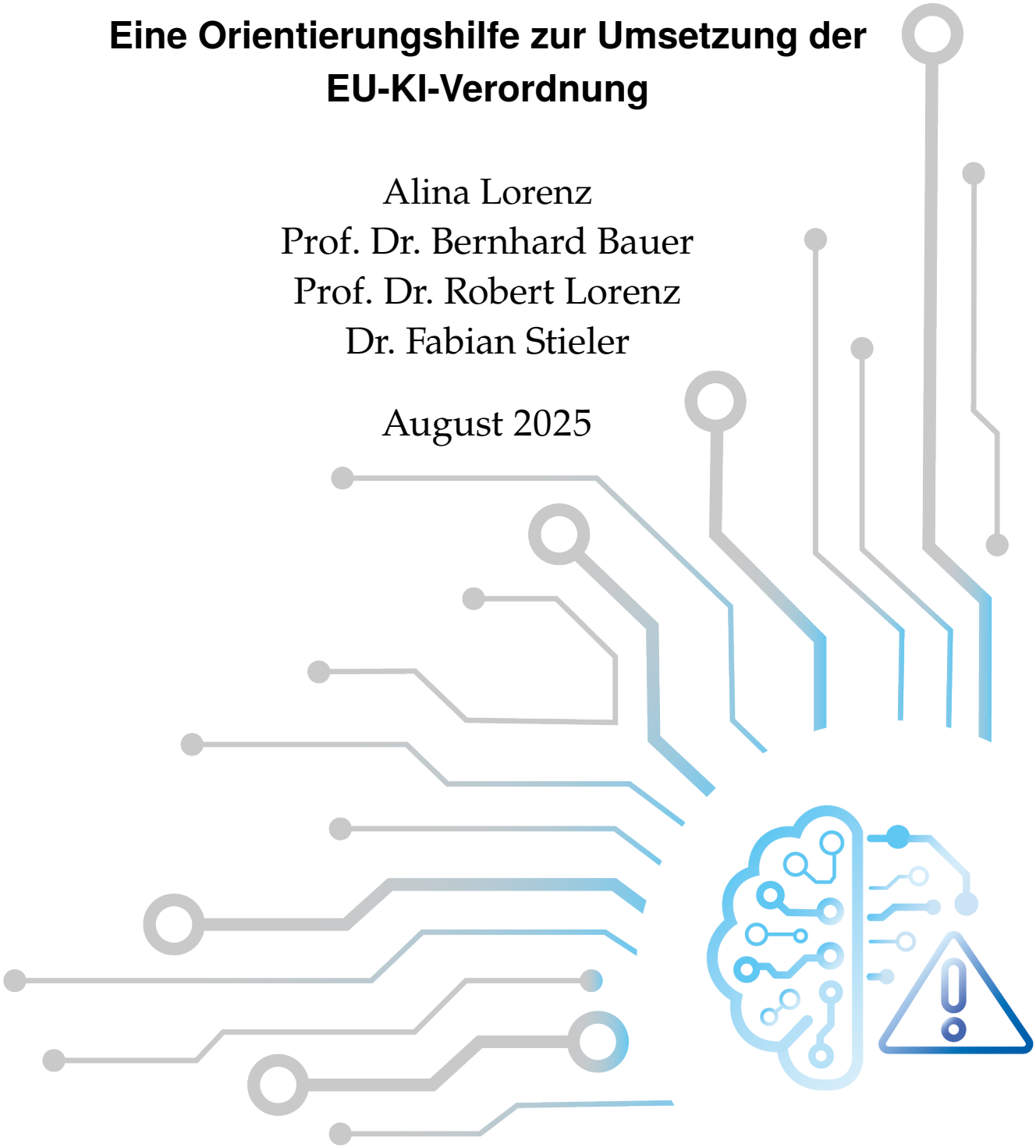
Audit- und Wissensplattform für Künstliche Intelligenz

KI-Risikomanagement

Eine Orientierungshilfe zur Umsetzung der EU-KI-Verordnung

Alina Lorenz
Prof. Dr. Bernhard Bauer
Prof. Dr. Robert Lorenz
Dr. Fabian Stieler

August 2025



Executive Summary

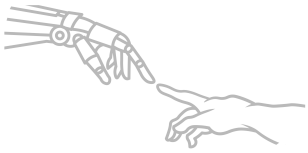
Die EU-KI-Verordnung (KI-VO) fordert für Hochrisiko-KI-Systeme die Einrichtung eines Risikomanagementsystems, das den kompletten Lebenszyklus eines KI-Systems von dessen Entwicklung bis zu dessen Betrieb begleitet.

Wir stellen ein strukturiertes Vorgehensmodell zur operativen Umsetzung eines integrierten Risikomanagements in Organisationen vor, das alle Vorgaben der KI-VO berücksichtigt.

Disclaimer
Die Lesbarkeit dieses Textes wurde mithilfe von KI überarbeitet. Zur besseren Lesbarkeit wurde das generische Maskulinum verwendet. Die verwendeten Personenbezeichnungen beziehen sich – sofern nicht anders kenntlich gemacht – auf alle Geschlechter.

Publisher
Audit- und Wissensplattform für vertrauenswürdige KI (AWIKI)
https://www.awiki.eu
August 2025

Inhaltsverzeichnis



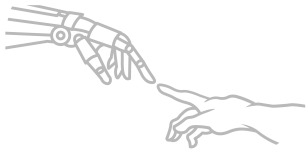
1 Einleitung

Die europäische KI-Verordnung (KI-VO) fordert von KI-Systemen mit hohem Risiko einen komplexen regulatorischen Rahmen, der darauf abzielt, systematisch relevante Risiken für Gesundheit, Sicherheit und Grundrechte zu ermitteln und zu mindern [1]. Zu diesem Zweck verlangt die KI-VO die Etablierung eines Risikomanagementsystems in Zusammenwirken mit einem Qualitätsmanagementsystem. Diese sollen gewährleisten, dass Anbieter und Betreiber von KI-Systemen geeignete Maßnahmen zur Abschwächung von bekannten und vernünftigerweise vorhersehbaren Risiken von KI-Systemen ergreifen. Hierbei werden mögliche Risiken, die sich aus der Wechselwirkung zwischen dem KI-System und dem Umfeld ergeben, inklusive der Berücksichtigung vorhersehbaren Missbrauchs des KI-Systems, eingeschlossen.

[1]: Das Europäische Parlament und der Rat der Europäischen Union (2024), *EU Artificial Intelligence Act*

Die operative Einführung und Umsetzung des geforderten Risikomanagementsystems erfordert die Etablierung, Integration, kontinuierliche Überprüfung, Aktualisierung und Dokumentation von organisatorischen Strukturen und eines Risikomanagementprozesses, der den gesamten Lebenszyklus des KI-Systems begleitet. Durch den Risikomanagementprozess werden systematisch und kontinuierlich konkrete Risiken ermittelt, bewertet und gemindert. Die Wahrnehmung dieser Aufgaben müssen unabhängige Risikomanagementteams mit passenden Kompetenzen verantworten. In dieser Publikation schlagen wir ein systematisches Vorgehen für die Schaffung der notwendigen Strukturen und Prozesse vor.

Dazu legen wir zuerst verschiedene Grundlagen für die operative Umsetzung des geforderten Risikomanagementsystems dar. Neben relevanten rechtlichen Vorgaben durch die KI-VO stellen wir einschlägige internationale Normen und Empfehlungen für den Umgang mit Risiken in Organisationen vor, die Unternehmen bei



der Umsetzung dieser Vorgaben unterstützen können. Zudem geben wir einen Überblick über verschiedene KI-spezifische Risikotaxonomien.

Danach erläutern wir im Detail ein systematisches Vorgehen bei der Einführung und Umsetzung eines Risikomanagementsystems, das mit der KI-VO konform ist und relevante internationale Normen, Empfehlungen und Best Practices berücksichtigt. Dazu gehört die Etablierung von Prinzipien des KI-Risikomanagements, die Durchführung verschiedener systematischer Schritte zur Einführung eines Risikomanagements und die operative Umsetzung eines kontinuierlichen und iterativen Risikozyklus zur Steuerung konkreter Risiken eines KI-Systems mit dessen Einbettung in dessen Produktlebenszyklus. Daran anschließend leiten wir die notwendigen Risikokompetenzen für die Besetzung von geeigneten Risikomanagementteams ab, die die geschilderte Einführung und Umsetzung eines Risikomanagementsystems verantworten.

Mit dem Reifegradmodell für das KI-Risikomanagement beschreiben wir ein systematisches Vorgehen, mit dem Organisationen feststellen können, in welchem Maß relevante Strukturen, Prozesse, Kompetenzen und kulturelle Voraussetzungen bereits vorhanden und wirksam sind – und wo noch gezielter Entwicklungsbedarf besteht.

Abschließend erörtern wir, inwieweit auch KI-Systeme mit begrenztem Risiko und Allzweck-KI-Systeme vom geschilderten Vorgehen zur Risikosteuerung profitieren können.

Regulatorische Einordnung und Normen

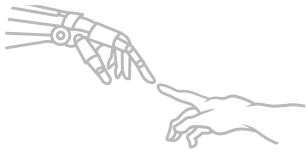
2

In diesem Kapitel erläutern wir die regulatorischen Grundlagen für ein systematisches Risikomanagement von KI-Systemen im Sinne der KI-VO. Wir konzentrieren uns dabei auf drei aufeinander bezogene Inhalte.

Erstens unterscheiden wir zwischen der verpflichtenden Risikoeinstufung eines KI-Systems gemäß KI-VO und der kontextbezogenen Bewertung konkreter Risiken eines KI-Systems. Die Risikoeinstufung legt fest, ob und in welchem Umfang ein KI-System unter die KI-VO fällt. Die KI-VO sieht dazu Regulierungen nach vier verschiedenen Stufen vor. Diese Stufen richten sich nach dem mit der Anwendung verbundenen Risiko für die Gesundheit, die Sicherheit oder die Grundrechte natürlicher Personen. Neben Anwendungen mit unannehmbaren Risiko (diese werden verboten) werden Kriterien für Anwendungen mit hohem, begrenztem und geringem Risiko definiert. Die Risikobewertung hingegen ist Bestandteil des betrieblichen Risikomanagements für KI-Systeme mit hohem Risiko (nach obiger Einstufung) und dient der vertieften Analyse und Behandlung konkreter Risiken im jeweiligen Anwendungskontext.

Zweitens fassen wir den rechtlichen Rahmen der KI-VO für das Risikomanagement von KI-Systemen mit hohem Risiko zusammen. Dabei wird deutlich, wie sich die regulatorischen Vorgaben in der Praxis systematisch, nachvollziehbar und rechtssicher umsetzen lassen.

Drittens beleuchten wir die Rolle von Normen und Standards als zentrale Instrumente zur technischen und organisatorischen Umsetzung gesetzlicher Anforderungen. Wir zeigen das Zusammenspiel zwischen Recht und Normung: Während das Recht verbindliche Schutzziele definiert, konkretisieren Normen diese durch praxistaugliche und überprüfbare Vorgaben. Die KI-VO verweist dabei ausdrücklich auf harmonisierte Normen und den Stand der Technik. Da die von der Europäischen Kommission beauftragten europäischen Normen



zur Umsetzung der KI-VO derzeit noch in Ausarbeitung sind, orientieren wir uns an bestehenden, international anerkannten Standards – etwa ISO/IEC 23894 (KI-Risikomanagement) oder dem NIST AI Risk Management Framework.

2.1 Risikoeinstufung versus Risikobewertung

Der risikobasierte Ansatz ist der zentrale Leitgedanke der KI-VO und strukturiert deren Systematik. Ziel ist es, Anforderungen an KI-Systeme nicht pauschal zu stellen, sondern differenziert in Abhängigkeit vom potenziellen Risiko. Diese Differenzierung wird durch zwei sich ergänzende Prozesse operationalisiert: die Risikoeinstufung und – bei entsprechender Kategorisierung – die Risikobewertung.

Risikoeinstufung

Legt fest, ob ein KI-System unter die KI-VO fällt, und bestimmt Art und Umfang der zu erfüllenden Anforderungen

Die **Risikoeinstufung** ist die formale Umsetzung dieses Grundgedankens. Sie ist für alle KI-Systeme verpflichtend durchzuführen und dient dazu, das System einer der in der Verordnung definierten Risikokategorien zuzuordnen. Die Einstufung ist erforderlich, um zu klären, ob und in welchem Umfang regulatorische Anforderungen greifen:

- Verbotene KI-Systeme (Art. 5).
- Hochrisiko-KI-Systeme (Art. 6 in Verbindung mit Anhang I und III): Solche Systeme unterliegen umfassenden Anforderungen, unter anderem zur Einrichtung eines Risikomanagements, dessen Umsetzung wir in diesem Leitfaden umfassend beschreiben.
- KI-Systeme mit begrenztem Risiko (Art. 50): Solche Systeme unterliegen bestimmten Transparenzpflichten.
- KI-Systeme mit geringem Risiko: Hierunter fallen alle KI-Syteme, die zu keine der vorherigen Kategorien gehören. Diese unterliegen keinerlei Anforderungen durch die KI-VO.

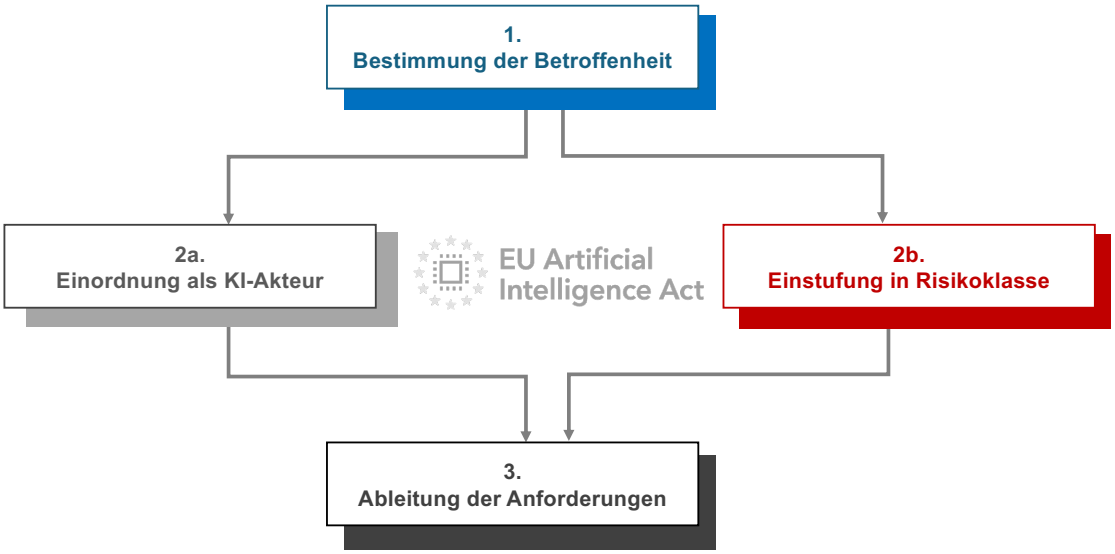


Abbildung 2.1: Risikoeinstufung in 4 Schritten.

Die Risikoeinstufung erfolgt in der Praxis häufig entlang eines vierstufigen Schemas, das sich als methodisch hilfreich erwiesen hat (siehe Abbildung 2.1):

- **Schritt 1**: Bestimmung der Betroffenheit – Wer oder was ist vom Einsatz des KI-Systems potenziell betroffen?
- **Schritt 2a**: Einordnung als KI-Akteur – Welche Rolle nimmt die Organisation ein (z.B. Anbieter, Betreiber)?
- **Schritt 2b**: Einstufung in Risikoklasse – Fällt das KI-System unter eine der Risikokategorien gemäß KI-VO?
- **Schritt 3**: Ableitung der Anforderungen – Welche regulatorischen Pflichten folgen daraus?

Wird ein System als Hochrisiko-KI eingestuft, sieht die KI-VO zusätzlich eine verpflichtende Risikobewertung vor. Diese ist Teil des umfassenden Risikomanagementsystems, das Anbieter solcher Systeme gemäß Art. 6 der Verordnung implementieren müssen. Ziel ist es, konkrete Risiken im Anwendungskontext systematisch zu identifizieren, zu bewerten und geeignete Maßnahmen zur Risikominderung zu treffen.

Die beiden Verfahren stehen in **direkter funktionaler**

Beziehung: Die Risikobewertung ist abhängig vom Ergebnis der vorangegangenen Einstufung. Nur wenn ein KI-System als hochriskant klassifiziert wurde, entsteht die Verpflichtung zur weiterführenden Bewertung und Kontrolle im betrieblichen Kontext.

Risikobewertung

Liefert eine fundierte, kontextbezogene Einschätzung konkreter Risiken eines KI-Systems und bildet die Grundlage für gezielte Maßnahmen zur Risikominderung

Eine **Risikobewertung** ist im Sinne eines iterativen Prozesses vorzunehmen. Dieser orientiert sich an etablierten Verfahren des betrieblichen Risikomanagements und umfasst die folgenden Schritte (siehe Abbildung Abbildung 2.2):

- Auswahl eines Risikoteams
- Situationsanalyse
- Risikoidentifikation
- Risikoanalyse
- Risikobewertung
- Risikobehandlung
- Wirkungsanalyse
- Dokumentation

In Kapitel 5 beschreiben wir jeden dieser Schritte im Gesamtzusammenhang im Detail.

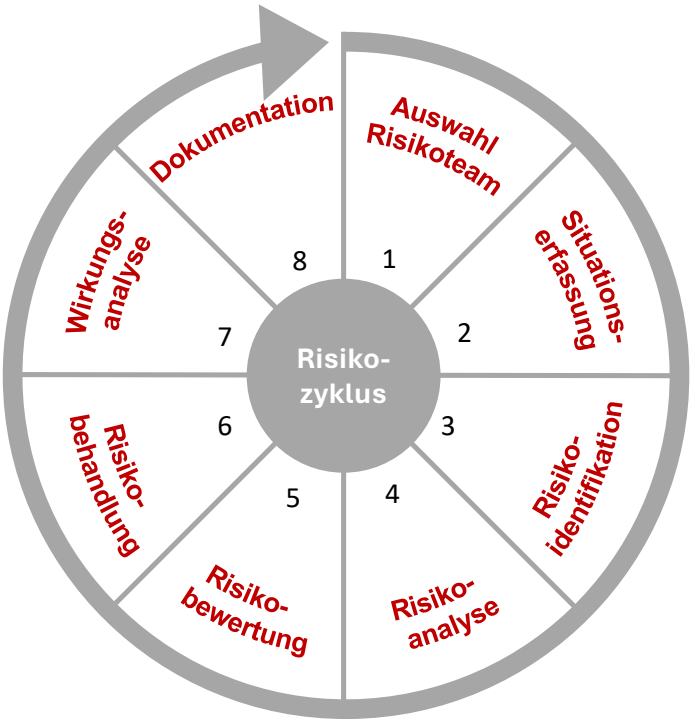


Abbildung 2.2: Risikobewertung in acht Schritten.

Risikoeinstufung und Risikobewertung sind zwei klar voneinander abgrenzbare, aber aufeinander aufbauende Verfahren im Rahmen der KI-VO. Während die Risikoeinstufung das „Ob“ und „Wie viel“ der regulatorischen Kontrolle bestimmt, beantwortet die Risikobewertung das „Wie konkret“ im Umgang mit identifizierten Risiken. Nur durch das Zusammenspiel beider Prozesse können die Schutzziele der Verordnung wirksam und praxisnah umgesetzt werden.

2.2 Rechtlicher Rahmen und Anforderungen der KI-VO

Die KI-VO setzt umfangreiche Anforderungen an das Risikomanagement für KI-Systeme mit hohem Risiko, um deren sicheren und verantwortungsvollen Einsatz zu gewährleisten. Tabelle 2.1 fasst die wesentlichen Kernforderungen der KI-VO zum Risikomanagement und ihre entsprechenden Quellen zusammen.

Zentrales Element ist ein risikobasierter Ansatz, der darauf abzielt, potenzielle Risiken entlang des gesamten Lebenszyklus eines KI-Systems zu identifizieren, zu bewerten und zu mindern (Artikel 9 Absatz 2). Dies umfasst nicht nur die bestimmungsgemäße Nutzung, sondern auch mögliche Risiken durch vernünftigerweise vorhersehbare Fehlanwendungen (Artikel 9 Absatz 2(b)). Zu den wichtigsten Vorgaben zählt die verpflichtende Einrichtung, Dokumentation und regelmäßige Aktualisierung eines umfassenden Risikomanagementsystems, das alle Phasen eines KI-Systems abdeckt – von der Entwicklung und Bereitstellung bis hin zum laufenden Betrieb (Artikel 9 Absatz 1 und Absatz 2). Regelmäßige Testverfahren sind notwendig, um die Wirksamkeit implementierter Maßnahmen zur Risikovermeidung und -minderung zu bestätigen (Artikel 9 Absätze 6 - 8). Im Hinblick auf den Schutz vulnerabler Gruppen, wie Minderjähriger, wird außerdem gefordert, dass das System besondere Rücksicht auf potenziell nachteilige Auswirkungen nimmt (Artikel 9 Absatz 9).

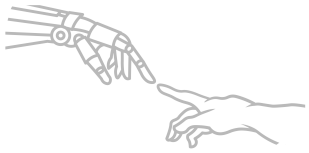


Tabelle 2.1: Forderungen der KI-VO zur Einrichtung eines Risikomanagementsystems für KI-Systeme mit hohem Risiko.

Forderung	Quelle	Beschreibung
Einrichtung Risikomanagement	Artikel 9 Absatz 1	Die KI-VO fordert die Einrichtung, Anwendung, Dokumentation und Aufrechterhaltung eines Risikomanagements.
Lebenszyklus-übergreifendes Risikomanagement	Artikel 9 Absatz 2	Das Risikomanagement muss während des gesamten Lebenszyklus eines KI-Systems angewendet werden, einschließlich Design, Entwicklung, Bereitstellung und Betrieb.
Überprüfung und Aktualisierung	Artikel 9 Absatz 2	Das Risikomanagementsystem muss regelmäßig systematisch überprüft und aktualisiert werden
Risikoidentifikation und -bewertung	Artikel 9 Absatz 2(a)-(c)	Risiken, die durch das KI-System entstehen können, insbesondere solche, die Gesundheit, Sicherheit und Grundrechte betreffen, sollen identifiziert und bewertet werden
Risiken bei Fehlanwendung	Artikel 9 Absatz 2(b)	Neben Risiken bei bestimmungsgemäßer Verwendung des KI-Systems müssen auch Risiken durch vernünftigerweise vorhersehbare Fehlanwendung betrachtet werden
Risikenvermeidung und -minderung	Artikel 9 Absätze 2(d), 4, 5	Nach der Risikobewertung müssen geeignete Maßnahmen zur Risikominderung implementiert werden, um sicherzustellen, dass identifizierte Risiken vermieden oder wirksam auf ein vertretbares Maß gemindert werden
Testverfahren	Artikel 9 Absätze 6-8	KI-Systeme müssen getestet werden, um die Wirksamkeit von Maßnahmen zur Risikominderung zu überprüfen
Schutzbedürftige Gruppen	Artikel 9 Absatz 9	Bei der Risikobewertung und -minderung müssen insbesondere potenziell nachteilige Auswirkungen auf schutzbedürftige Gruppen (wie Minderjährige) berücksichtigt werden
Dokumentation	Artikel 11 Anhang IV	Das Risikomanagementsystem und die Ergebnisse und Quellen vorhersehbarer Risiken müssen dokumentiert und beschrieben werden
Transparenz und Nutzerinformationen	Artikel 9 Absatz 5(c) Artikel 13	KI-Systeme müssen transparent konzipiert sein und Informationen über bekannte und vorhersehbare Umstände, die zu Risiken führen, bereitstellen

Die KI-VO schreibt zudem vor, dass alle erkannten Risiken systematisch dokumentiert und Informationen über bekannte und vorhersehbare Risiken transparent für die Nutzer zur Verfügung gestellt werden müssen (Artikel 11, Anhang IV, Artikel 9 Absatz 5(c) und Artikel 13). Diese umfassende Dokumentation soll sicherstellen, dass ein Höchstmaß an Transparenz besteht und Nutzer die potenziellen Risiken verstehen können. Die Anfor-

derungen an Transparenz, Schutz und systematische Risikoüberwachung bilden somit das Fundament für die Entwicklung eines rechtskonformen und robusten Risikomanagementsystems im Sinne der KI-VO.

2.3 Normen als Umsetzungsinstrumente gesetzlicher Anforderungen

Um die geschilderten Anforderungen technisch und organisatorisch umsetzbar zu machen, verweist die KI-VO explizit auf harmonisierte Normen und den Stand der Technik. Gesetzgebung und Normung wirken dabei komplementär: Während das Recht die Ziele und Pflichten festlegt, liefern Normen die konkreten Mittel zu deren Umsetzung. Diese Normen werden von anerkannten Organisationen wie ISO, IEC, CEN/CENELEC oder ETSI entwickelt und ermöglichen es, regulatorische Vorgaben in prüfbare Verfahren, Kriterien und Dokumentationsanforderungen zu übersetzen – etwa für das Risikomanagement, die Konformitätsbewertung oder die technische Dokumentation. Die Wechselwirkung von Gesetzgebung und Normung im Rahmen der KI-VO ist in Abbildung 2.3 dargestellt.

Die Normung übernimmt somit eine Brückenfunktion: Sie macht abstrakte Rechtsvorgaben technisch konkret, nachvollziehbar und anwendbar. Normen sind dabei kein Ersatz für das Recht, sondern ein zentrales Werkzeug für dessen Umsetzung in der Praxis. Werden bei Entwicklung und Einsatz eines KI-Systems harmonisierte Normen angewendet, geht der Gesetzgeber zudem automatisch davon aus, dass das KI-System mit den KI-VO-Anforderungen, welche die Normen abdecken, konform ist. Man spricht von der sogeannnten „Konformitätsvermutung" – einem zentralen Prinzip im europäischen Binnenmarkt.

Für die Umsetzung der vielfältigen Anforderungen an das Risikomanagement können verschiedene etablierte

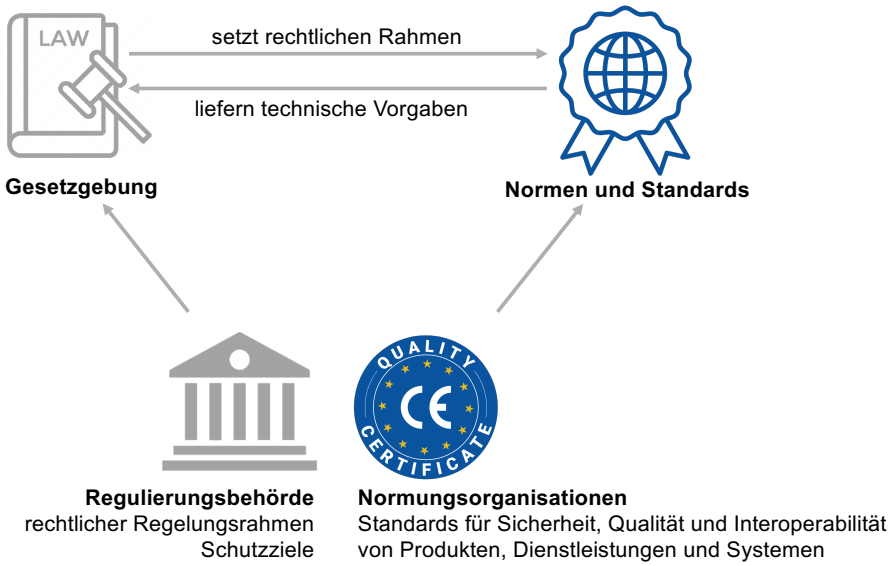


Abbildung 2.3: Wechselwirkung zwischen Gesetzgebung und Normung.

internationale Normen, Standards und Empfehlungen herangezogen werden. Zwei der wichtigsten stammen von der Internationalen Organisation für Normung (ISO) und der Internationalen Elektrotechnischen Kommission (IEC), die gemeinsam an harmonisierten Normen für KI und Risikomanagement arbeiten. Ergänzend bieten das National Institute of Standards and Technology (NIST) der USA und das Institute of Electrical and Electronics Engineers (IEEE) anerkannte Leitlinien, die vielfach Anwendung finden – insbesondere solange europäische Normen zur KI-VO noch in Ausarbeitung sind.

[2]: NIST (2024), *Artificial Intelligence Risk Management Framework*

Das Rahmenwerk NIST Artificial Intelligence Risk Management Framework (AI RMF) [2] definiert Risiken von KI-Systemen und empfiehlt Risikomanagement-Aufgaben für die Entwicklung und den Betrieb vertrauenswürdiger KI-Systeme.

[3]: IEEE (2024), *Ethically Aligned Design*

Das umfassende multidisziplinäre Rahmenwerk IEEE Ethically Aligned Design (EAD) [3] wurde in globaler Kooperation von Experten aus verschiedenen Bereichen wie Ethik, Recht, Sozialwissenschaften, Philosophie und verschiedenen Technologiegebieten entwickelt. Es besteht aus einer Reihe von ethischen Prinzipien

und Empfehlungen für deren Umsetzung, die die ethische Entwicklung von autonomen und intelligenten Systemen leiten sollen, so dass dabei das menschliche Wohlergehen in den Vordergrund gestellt wird, sowie Menschenrechte und ethische Standards eingehalten werden.

Tabelle 2.2: Schwerpunkte international relevanter Normen und Empfehlungen für Risikomanagement.

Norm	Schwerpunkte
NIST AI RMF Rahmenwerk für das Risikomanagement vertrauenswürdiger KI-Systeme	**Grundlegende Definitionen** ► Risiken von KI-Systemen ► Vertrauenswürdige KI-Systeme **Risikomanagementaufgaben** ► Etablierung von Governance-Strukturen ► Identifikation von Risiken (soziotechnischer Kontext) ► Bewertung, Analyse und Überwachung von Risiken ► Priorisierung und Behandlung von Risiken
IEEE EAD Rahmenwerk für ethische Prinzipien und deren Umsetzung in KI-Systemen	**Prinzipien** ► Schutz der Menschenrechte ► Vorrang menschlichen Wohlergehens ► Verantwortlichkeit und Rechenschaftspflicht ► Förderung von Transparenz ► Bewusstsein für Risiken durch Fehlanwendung ► Sicherstellung von Datenhoheit ► Kompetenz für den sicheren und effektiven Einsatz **Umsetzungsaspekte** ► Einbettung von Werten ► Policies ► Recht ► Designmethoden ► Persönliche Daten und individuelle Autorität ► Ethik in KI-Systemen ► Menschliches Wohlergehen

Die Norm ISO 31000:2018 [4] definiert allgemeine Leitlinien für den Umgang mit Risiken in einer Organisation. Diese Leitlinien stellen Empfehlungen dar, die auf Organisationen spezifisch angepasst werden können. Die Norm kann für jede Art von Organisation und Risiken über die gesamte Lebensdauer der Organisation und auf allen Organisationsebenen angewendet werden.

[4]: DINMedia (2018), *DIN ISO 31000: 2018-10: Risikomanagement - Leitlinien*

Die Norm ISO/IEC 23894:2023 [5] basiert auf ISO 31000, postuliert dieselben allgemeinen Leitlinien, Komponenten und Schwerpunkte eines Risikomanagementsystems, und enthält spezifische Anleitungen, Anpassungen und Ergänzungen für deren Anwendung auf KI-Systeme.

[5]: ISO (2023), *ISO 23894: 2023: Information technology, Artificial intelligence, Guidance on risk management*

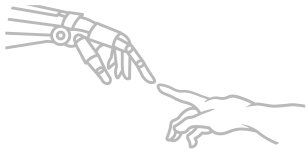
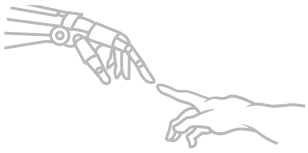


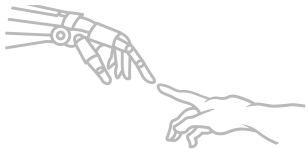
Tabelle 2.3: Schwerpunkte international relevanter Normen und Empfehlungen für Risikomanagement (Fortsetzung).

Norm	Schwerpunkte
ISO 31000:2018 Allgemeine Norm für Risikomanagement	**Risikomanagement-Prinzipien** Risikomanagement ► bezweckt Wertschöpfung ► ist in die Organisation integriert ► ist strukturiert und umfassend ► ist an den Anwendungskontext angepasst ► ist ein dynamischer und iterativer Prozess ► basiert auf Informationen und Erwartungen ► ist abhängig von kulturellen und menschlichen Faktoren ► bezieht beteiligte Akteure ein ► wird kontinuierlich verbessert **Risikomanagement-Rahmenwerk** ► Etablierung durch die Organisationsleitung ► Integration in Organisationsziele und -prozesse ► Gestaltung und Implementierung ► Regelmäßige Bewertung ► Kontinuierliche Verbesserung **Risikomanagementzyklus** ► Kommunikation und Konsultation mit Akteuren ► Festlegung von Kontext und Risikokriterien ► Risikoidentifikation, -analyse und -bewertung ► Risikosteuerung und -mitigation ► Risikoüberwachung ► Risikoberichterstattung
ISO/IEC 23894:2023 Norm für das Risikomanagement von KI-Systemen	**Risikomanagement-Prinzipien** Wie ISO 31000 mit ergänzenden spezifischen Folgerungen, die sich aus den Prinzipien für die Entwicklung und den Einsatz von KI-Systemen ergeben **Risikomanagement-Rahmenwerk** Wie ISO 31000 mit ergänzenden spezifischen Anleitungen für die Umsetzung des Rahmenwerks bei Entwicklung und Einsatz von KI-Systemen **Risikomanagementzyklus** Wie ISO 31000 mit ergänzenden und angepassten spezifischen Anleitungen für die Umsetzung der Prozessschritte bei Entwicklung und Einsatz von KI-Systemen **KI-Lebenszyklus** Abbildung des Risikomanagementzyklus auf den KI-Lebenszyklus

Obwohl diese Normen und Empfehlungen unterschiedliche Ansätze wählen, betonen sie alle, dass ein wirksames Risikomanagementsystem umfassend, gut integriert und proaktiv sein muss, sowie eine Aufgabe der Leitung einer Organisation ist. Es sollte über den gesamten KI-Lebenszyklus hinweg systematisch Risiken identifizieren, kontinuierlich überwachen, ethische und



rechtliche Aspekte berücksichtigen und alle relevanten Akteure einbeziehen. Die Tabellen 2.2 und 2.3 fassen wesentliche Komponenten und Schwerpunkte der genannten Normen und Empfehlungen zusammen.



3 KI-Risikotaxonomien

Die in der KI-VO verwendete Risikodefinition entspricht der klassischen und allgemeinen Risikodefinition, wie sie auch in anderen Bereichen und Disziplinen angewendet wird.[1] Die KI-VO betont jedoch auch, dass das KI-Risiko vielfältige Ursachen und Auswirkungen haben kann, die stark vom spezifischen Kontext der Entwicklung und des Einsatzes von KI abhängen. Damit geht sie über die traditionelle Risikobetrachtung hinaus, indem sie die Multidimensionalität von KI-Risiken berücksichtigt. Dies ist notwendig, um das volle Spektrum potenzieller Auswirkungen zu erfassen, denn KI-Risiken können sich sowohl auf individuelle Rechte und Interessen als auch auf gesellschaftliche, wirtschaftliche und politische Strukturen und Prozesse auswirken.

1: **KI-VO Art. 3 Abs. 2**: Risiko ist die Kombination aus der Wahrscheinlichkeit des Eintritts eines Schadens und der Schwere dieses Schadens

Ein im Jahr 2024 vom Massachusetts Institute of Technology (MIT) öffentlich bereitgestellter Katalog von KI-Risiken verdeutlicht die Vielfalt von KI-Risiken.[2] Dieser Katalog basiert auf einer systematischen Analyse existierender Studien. Er enthielt bei seinem Start etwas über 700 untersuchte KI-Risiken und ist mittlerweile (Stand Juni 2025) auf über 1600 verschiedene bekannte KI-Risiken angewachsen. Um diese Risiken handhabbar und verständlich zu machen, wurden sie bzgl. verschiedener Taxonomien in Klassen unterteilt. Hierfür wurden verschiedene existierende Risiko-Taxonomien untersucht und unter anderem die folgenden im KI-Kontext besonders relevanten Taxonomien näher betrachtet:

2: https://airisk.mit.edu

- Taxonomie nach Domänen
- Taxonomie nach Intention
- Taxonomie nach Verursacher
- Taxonomie nach Zeitpunkt des Entstehens

Abbildung 3.1 zeigt exemplarisch das Ergebnis der Klassifikation nach Domänen. Jeder Balken entspricht einer domänenspezifischen Risikoklasse und ist unterteilt in weitere Unterklassen. Für jede Klasse ist angegeben,

welcher Anteil der Menge aller Risiken in diese Klasse fallen. Beispielsweise entfallen 4% aller Risiken auf den Bereich der Fehlinformation.

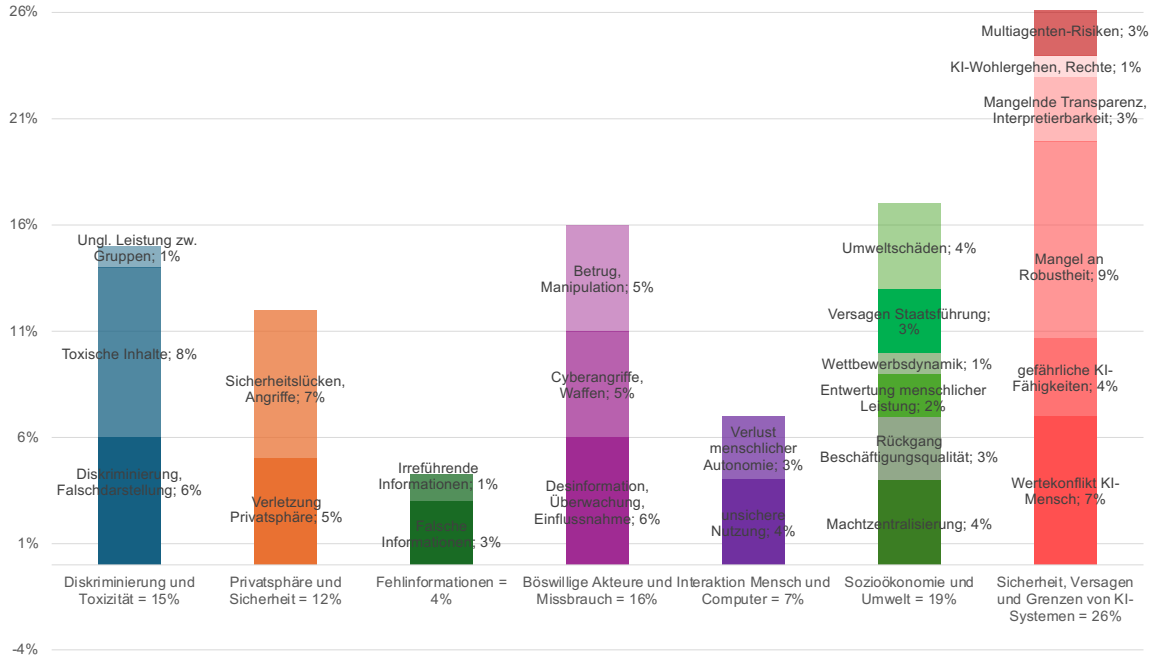
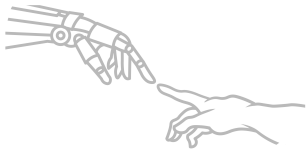


Abbildung 3.1: KI-Risiken klassifiziert nach Domänen nach `https://airisk.mit.edu` (Juni 2025).

Die in Abbildung 3.1 in komprimierter Darstellung erfassten KI-Risiken zeigen die Vielfalt und Komplexität der KI-Risiken. Risiko-Taxonomien sind dabei ein nützliches Werkzeug, Risiken systematisch zu identifizieren und ihre Bedeutung sowie Dringlichkeit zu priorisieren. Für das KI-Risikomanagement sind gut gewählte und definierte Taxonomien besonders nützlich, da sie Folgendes unterstützen:

- **Bewertung von Risiken**: Jede Risikoklasse kann spezifische Risikobewertungmethoden erfordern.
- **Standardisierung von Methoden und Maßnahmen**: Durch Klassifizierung können Standardmethoden und -Maßnahmen für ähnliche Risikotypen entwickelt werden.
- **Gezielte Ressourcenzuweisung**: Ein strukturiertes System hilft Entscheidungsträgern, Risiken nach ihrer Schwere und Wahrscheinlichkeit zu bewerten und Ressourcen gezielt für die Bewältigung der dringlichsten Risiken einzusetzen.
- **Klärung von Verantwortlichkeiten**: Verschiedene Risikoklassen können verschiedenen Abteilungen



oder Akteuren zugewiesen werden, die für deren Überwachung und Minderung zuständig sind.

Taxonomien bieten also ein umfassendes Rahmenwerk zur strukturierten Verwaltung von KI-Risiken. Sie unterstützen das Risikomanagement bei der Identifizierung, Bewertung, Priorisierung und Minderung von Risiken, was zu einem effizienteren und robusteren Risikomanagement führt. Im Folgenden werden relevante Taxonomien für KI-Risiken beschrieben und ihre Bedeutung für das KI-Risikomanagement diskutiert. Dabei betrachten wir nicht nur die bereits erwähnten Taxonomien, die im MIT-Katalog unterschieden wurden, sondern folgende zusätzliche aus unserer Sicht für die weitere Diskussion relevante Taxonomien:

- Taxonomie nach Systemrelevanz (in der KI-VO verwendet)
- Taxonomie nach Themen- und Handlungsfeldern (von den Autoren aus der Praxis heraus entwickelt)
- Taxonomie nach Produktlebenszyklusphase (von den Autoren aus der Praxis heraus entwickelt)

Risikotaxonomie nach Zeitpunkt des Entstehens (MIT-Katalog)

Die **Risikotaxonomie nach Zeitpunkt des Entstehens** betrachtet Risiken in Bezug auf den Zeitpunkt ihres Auftretens (wann tritt ein bestimmtes Risiko auf?). Dabei können unter anderem folgende Zeitpunkte unterschieden werden:

- **Risiken vor der Bereitstellung**: Risiken, die während der Entwicklungsphase eines KI-Systems entstehen, etwa durch Designmängel, fehlerhafte Datensätze oder unzureichende Tests.
- **Risiken nach der Bereitstellung**: Risiken, die auftreten, wenn ein KI-System in der realen Welt eingesetzt wird, wie unvorhergesehene Fehlfunktionen oder neue Sicherheitslücken.

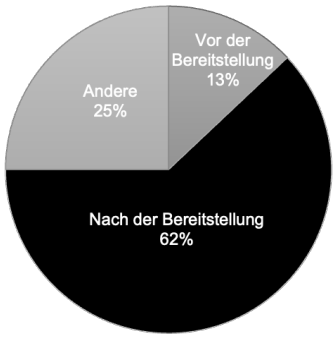


Abbildung 3.2: KI-Risiken klassifiziert nach Zeitpunkt des Entstehens nach https://airisk.mit.edu (Juni 2025).

Abbildung 3.2 veranschaulicht die Anteile der vor und nach der Bereitstellung auftretenden Risiken an der Menge aller Risiken aus dem MIT-Katalog. Hierbei können 25% der Risiken keiner dieser beiden Klassen zugeordnet werden.

Diese Klassifizierung hilft, Risiken präventiv anzugehen und ermöglicht eine fortlaufende Überwachung nach der Bereitstellung von KI-Systemen. In der Entwicklungsphase liegt der Schwerpunkt auf gründlichen Tests, während nach der Einführung eine kontinuierliche Überwachung und Aktualisierung der Systeme erforderlich ist.

Risikotaxonomie nach Intention (MIT-Katalog)

Die **Risikotaxonomie nach Intention** unterscheidet, ob ein Risiko absichtlich oder unbeabsichtigt verursacht wird (wie tritt ein bestimmtes Risiko auf?):

- **Absichtlich verursachte Risiken**: Missbrauch von KI durch böswillige Akteure (z.B. Cyberangriffe oder absichtliche Manipulationen) oder in Kauf genommene negative Seiteneffekte.
- **Unbeabsichtigt verursachte Risiken**: Unvorhergesehene Konsequenzen durch Designfehler, Fehlfunktionen oder unerwartete Interaktionen mit KI-Systemen.

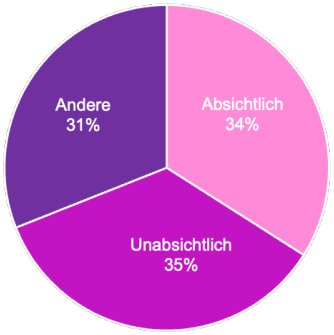


Abbildung 3.3: KI-Risiken klassifiziert nach Intention gemäß https://airisk.mit.edu (Juni 2025).

Abbildung 3.3 zeigt die Anteile der absichtlich und unabsichtlich herbeigeführten Risiken an der Menge aller Risiken aus dem MIT-Katalog.

Die Absicht bei der Verursachung eines Risikos ist wichtig, um die richtige Strategie zur Risikominimierung zu wählen. Absichtliche Risiken erfordern Sicherheitsmaßnahmen und Schutzvorkehrungen, während unbeabsichtigte Risiken durch Tests, Monitoring und Anpassungen des Systems im Laufe der Zeit minimiert werden können.

Risikotaxonomie nach Verursacher (MIT-Katalog)

Die **Risikotaxonomie nach Verursacher** klassifiziert Risiken basierend darauf, ob die Risiken durch Entscheidungen oder Handlungen eines Menschen oder der KI-Anwendung verursacht werden (warum tritt ein bestimmtes Risiko auf?):

- **Verursacht durch Menschen**: Schäden, die durch menschliche Handlungen (z.B. Angriffe, Fehler oder Nachlässigkeit) verursacht werden.

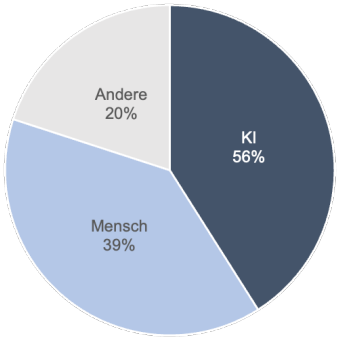


Abbildung 3.4: KI-Risiken klassifiziert nach Verursacher gemäß `https://airisk.mit.edu` (Juni 2025).

- **Verursacht durch KI**: Schäden, die direkt durch das Verhalten oder die Entscheidungen von KI-Systemen (z.B. Diskriminierung oder Datenschutzverletzungen) entstehen.

Abbildung 3.4 illustriert die Anteile der durch das KI-System und durch den Menschen verursachten Risiken an der Menge aller Risiken aus dem MIT-Katalog.

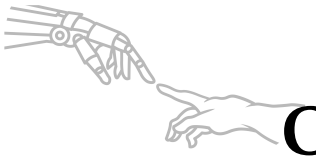
Die Identifizierung der Ursache eines Risikos kann für die Entwicklung gezielter Maßnahmen genutzt werden. Risiken, die auf menschliches Fehlverhalten zurückzuführen sind, erfordern Schulungen und Verhaltensrichtlinien, während technische Fehler durch Training und Tests behoben werden können.

Risikotaxonomie nach Domänen (MIT-Katalog)

Die **Risikotaxonomie nach Domänen** kategorisiert KI-Risiken nach dem Einsatzbereich oder der Domäne, in der sie auftreten. Zu den wichtigsten Domänen gehören:

- **Diskriminierung und Toxizität**: Verzerrungen in der Entscheidungsfindung des KI-Systems.
- **Datenschutz und Sicherheit**: Datenschutzverletzungen und Sicherheitslücken.
- **Mensch-Computer-Interaktion**: Verlust der Autonomie oder zu starkes Vertrauen in KI.
- **Böswillige Akteure und Missbrauch**: Gezielte Manipulation von KI-Systemen.
- **Sozioökonomische und ökologische Schäden**: Arbeitsplatzverluste, soziale Ungerechtigkeit, Umweltschäden.
- **Fehlinformationen**: Verbreitung falscher Informationen durch KI-Systeme.

Diese Klassifikation wurde bereits in Abbildung 3.1 illustriert und ermöglicht es, die Risiken in den jeweiligen Kontexten, in denen KI eingesetzt wird, gezielt zu adressieren. Beispielsweise können Risiken der Mensch-Maschine-Interaktion durch Maßnahmen zur Weiterbildung und Förderung der Transparenz gemindert werden, während Diskriminierungs-Risiken durch Training und Tests verhindert werden können.



Operative Umsetzung KI-Risikomanagement

4

Die KI-VO fordert für Hochrisiko-KI-Systeme die Einführung eines Risikomanagementsystems, das den gesamten Lebenszyklus eines Systems abdeckt – von der Entwicklung über den Betrieb bis zur Wartung. Ein solches System muss rechtliche, ethische, technische und organisatorische Anforderungen integrieren und kontinuierlich weiterentwickelt werden.

Ein strukturiertes und gelebtes (KI-)Risikomanagement ist in einem Unternehmen nicht per se vorhanden, sondern muss zuerst eingeführt werden. Wir beschreiben in diesem Kapitel die unterschiedlichen dafür notwendigen Schritte und Verantwortlichkeiten, damit die Anforderungen der KI-VO erfüllt werden. Basierend auf diesen Anforderungen und den genannten relevanten Normen und Empfehlungen identifizieren wir die folgenden Aufgabenbereiche für eine operative Umsetzung eines KI-Risikomanagements:

- Ausrichtung des KI-Risikomanagement an Prinzipien (Grundsätzen).
- Etablierung und Integration eines Risikomanagementprozesses mit kontinuierlicher Überprüfung und Anpassung.
- Etablierung eines Risikozyklus zur Risikosteuerung konkreter KI-Systeme.

Abbildung 4.1 gibt einen Überblick über das von uns vorgeschlagene Vorgehen und den Zusammenhang zwischen diesen Aufgabenbereichen. Risikomanagement ist nicht als isolierter Prozess zu betrachten, sondern als ein wesentlicher Bestandteil aller relevanten organisatorischen Aktivitäten und Entscheidungsprozesse. Deshalb ist die Einführung von Prinzipien ein wichtiger Baustein (linker Teil in Abbildung 4.1). Diese sind ein wesentlicher Teil der Befähigungs- und Bewusstseinsbildung, die den gesamten KI-Risikomanagementprozess begleiten muss. Zudem leiten sie die Etablierung des übergreifenden Risikomanagementprozesses durch die

Initiierung von Schulungsprogrammen und Trainingsmaßnahmen, um das Bewusstsein und das Verständnis für Risikomanagementpraktiken zu fördern.

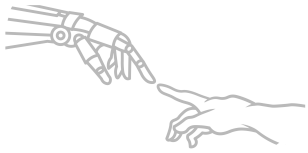
Start von Risikoschulungen

4.3 Zwischenfazit

Das geschilderte Vorgehen erfüllt zentrale Vorgaben der KI-Verordnung zur Einrichtung eines Risikomanagements. Tabelle 4.5 zeigt eine Übersicht über die bisher abgedeckten Vorgaben. Die bisher noch nicht abgedeckten Vorgaben betreffen den Risikozyklus (Kapitel 5) und sind dort berücksichtigt.

Tabelle 4.5: Berücksichtigung von Vorgaben der KI-VO (Zwischenfazit).

Forderung	**Quelle**	**Umsetzungsschritt**
Einrichtung Risikomanagement	Artikel 9 Absatz 1	Beschrieben in Kapitel 4.2 (Schritte 1 – 5)
Lebenszyklus-übergreifendes Risikomanagement	Artikel 9 Absatz 2	Beschrieben in Kapitel 4.2 (Schritt 3)
Überprüfung und Aktualisierung	Artikel 9 Absatz 2	Beschrieben in Kapitel 4.2 (Schritte 2-4) Teil des Risikozyklus (Kapitel 5)
Risikoidentifikation und -bewertung	Artikel 9 Absatz 2(a)-(c)	Teil des Risikozyklus (Kapitel 5)
Risiken bei Fehlanwendung	Artikel 9 Absatz 2(b)	Teil des Risikozyklus (Kapitel 5)
Risikenvermeidung und -minderung	Artikel 9 Absätze 2(d), 4, 5	Teil des Risikozyklus (Kapitel 5)
Testverfahren	Artikel 9 Absätze 6-8	Teil des Risikozyklus (Kapitel 5)
Schutzbedürftige Gruppen	Artikel 9 Absatz 9	Teil des Risikozyklus (Kapitel 5)
Dokumentation	Artikel 11 Anhang IV	Beschrieben in Kapitel 4.2 (Schritte 2-4) Teil des Risikozyklus (Kapitel 5)
Transparenz und Nutzerinformationen	Artikel 9 Absatz 5(c) Artikel 13	Beschrieben in Kapitel 4.2 (Schritte 2-4)



5 KI-Risikozyklus

Ein Teilschritt der in Kapitel 4.2 besprochenen Etablierung eines Risikomanagementprozesses ist die Planung eines Risikozyklus mit spezifischen Schritten für die Steuerung von und den Umgang mit Risiken einer konkreten KI-Anwendung inklusive Zuordnung von Rollen und Verantwortlichkeiten zu diesen Schritten. Entsprechend der Anforderungen der KI-VO und der zitierten Normen schlagen wir den acht Schritte umfassenden Risikozyklus aus Abbildung 5.1 vor.

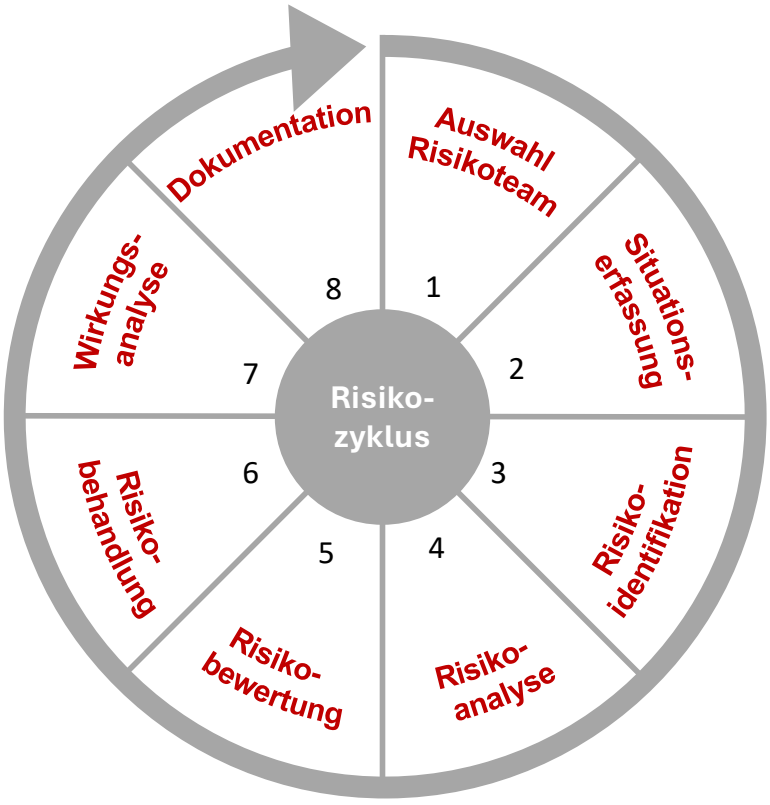


Abbildung 5.1: Schritte des Risikozyklus.

Die KI-VO und einschlägige Normen und Empfehlungen betonen die Notwendigkeit, das Risikomanagement über den gesamten Lebenszyklus eines KI-Systems hinweg zu integrieren, von der Konzeption über die Entwicklung und den Betrieb bis hin zur kontinuierlichen Überwachung und Anpassung. Entsprechend sollte der Risikozyklus wiederholt und zu festgelegten Zeitpunkten

entlang des Produktlebenszyklus der KI-Anwendung durchgeführt werden. In Schritt 3 in 4.2 haben wir dazu als Zeitpunkte für die Durchführung empfohlen:

- **Initiale Durchführung** nach der Konzeptionsphase.
- **Finale Durchführung** direkt vor der Produktivsetzungsphase.
- **Regelmäßige Durchführung während des Betriebs** in Re-Evaluierungs- und Wartungsphasen der KI-Anwendung.

Im Folgenden beschreiben wir die acht Schritte zur Durchführung des Risikozyklus mit ihren Aufgaben, Methoden und jeweiligen Ergebnissen, und begründen die Konformität des so gebildeten Risikozyklus mit den Anforderungen der KI-VO gemäß Kapitel 4.3. Die acht Schritte hängen dabei in ihrer Ausgestaltung vom Zeitpunkt der Durchführung des Risikozyklus, also von den jeweils betrachteten Produktlebenszyklusphasen, ab.

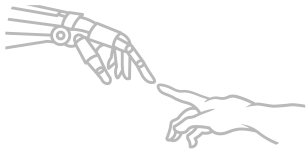
5.1 Risikobewertendenteam und Bewertungsrahmen

Risikobewertendenteam

Zuständig für die Durchführung des Risikozyklus für eine spezifische KI-Anwendung

Der erste Schritt im Risikozyklus einer KI-Anwendung beginnt mit der Bildung eines geeigneten Risikobewertendenteams. Damit das Team Risiken glaubwürdig und wirksam bewerten und steuern kann, muss es bestimmte Kriterien erfüllen:

- **Unabhängigkeit**: Die Teammitglieder dürfen nicht durch eigene wirtschaftliche oder organisatorische Interessen befangen sein. Zudem sollte die Leitung mit ausreichender Entscheidungs- und Steuerungskompetenz ausgestattet sein.
- **Interdisziplinarität:** Technische, rechtliche, ethische und betroffenenbezogene Perspektiven müssen gleichberechtigt berücksichtigt werden.



der Vorfälle gemäß den Anforderungen der KI-VO.

Ergebnis

Die Dokumentation fungiert damit als Brücke zwischen fachlicher Steuerung, operativer Umsetzung und regulatorischer Absicherung. Das Ergebnis dieses Schritts ist eine vollständige und fortlaufend gepflegte Dokumentation des Risikozyklus. Sie erlaubt eine konsistente Rückverfolgbarkeit aller Entscheidungen, Maßnahmen und Ergebnisse – und ist damit Grundlage für interne Qualitätssicherung ebenso wie für externe Prüfung, Zertifizierung oder regulatorische Kontrolle. Dazu fließt sie in die Dokumentation des gesamten Risikomanagementprozesses ein und dient so als Grundlage für dessen regelmäßige Überprüfung und Aktualisierung.

5.9 KI-VO Konformität des Risikomanagements

Der gesamte Risikomanagementprozess inklusive Risikozyklus erfüllt zusammengenommen alle zentralen Vorgaben der KI-VO zur Einrichtung eines Risikomanagements. Tabelle 5.6 zeigt eine Übersicht darüber, welche der Vorgaben in welchem Umsetzungsschritt berücksichtigt sind. Zwei Forderungen der KI-VO sind dabei implizit im Risikozyklus berücksichtigt:

- Die Betrachtung von Risiken bei Fehlanwendung
- und von Risiken für schutzbedürftige Gruppen.

Diese beiden durch die KI-VO besonders hervorgehobenen Arten von Risiken werden vom ersten Schritt des Risikozyklus an in den Checklisten mitberücksichtigt und sind danach automatisch Teil des kompletten Risikozyklus. Zudem erfüllt das Vorgehen zusätzliche zentrale Forderungen der einschlägigen internationalen Normen und Empfehlungen:

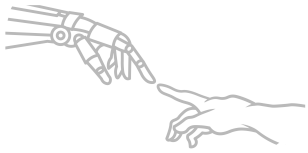
- Risikomanagement ist Führungsaufgabe: Kapitel 4.2 (Schritt 1)

- Risikomanagement bezieht alle beteiligten Akteure mit ein und erfordert die Festlegung von Verantwortlichkeiten: Kapitel 4.2 (Schritt 2), Kapitel 5.1 und Kapitel 5.6
- Risikomanagement wird nicht als isolierter Prozess verstanden, sondern als integraler Bestandteil der Unternehmensstrategie und -kultur und ist Teil aller relevanten organisatorischen Aktivitäten: Kapitel 4.2 (Schritte 2 und 3)
- Risikomanagement ist strukturiert, umfassend und an den Anwendungskontext angepasst: Kapitel 4.2 (Schritte 1 – 5), Kapitel 5.1 - 5.7

Tabelle 5.6: Berücksichtigung von Vorgaben der KI-VO.

Forderung	Quelle	Umsetzungsschritt
Einrichtung Risikomanagement	Artikel 9 Absatz 1	Beschrieben in Kapitel 4.2 (Schritte 1 – 5)
Lebenszyklus-übergreifendes Risikomanagement	Artikel 9 Absatz 2	Beschrieben in Kapitel 4.2 (Schritt 3)
Überprüfung und Aktualisierung	Artikel 9 Absatz 2	Beschrieben in Kapiteln 4.2 (Schritte 2-4) und 5.7
Risikoidentifikation und -bewertung	Artikel 9 Absatz 2(a)-(c)	Beschrieben in Kapiteln 5.3 - 5.5
Risiken bei Fehlanwendung	Artikel 9 Absatz 2(b)	Implizit berücksichtigt in Kapiteln 5.1 - 5.7
Risikenvermeidung und -minderung	Artikel 9 Absätze 2(d), 4, 5	Beschrieben in Kapitel 5.6
Testverfahren	Artikel 9 Absatz 6-8	Beschrieben in Kapitel 5.7
Schutzbedürftige Gruppen	Artikel 9 Absatz 9	Implizit berücksichtigt in Kapiteln 5.1 - 5.7
Dokumentation	Artikel 11 Anhang IV	Beschrieben in Kapiteln 4.2 (Schritte 2-4) und 5.8
Transparenz und Nutzerinformationen	Artikel 9 Absatz 5(c) Artikel 13	Beschrieben in Kapiteln 4.2 (Schritte 2-4) und 5.8

Um die Wirksamkeit der von uns empfohlenen Maßnahmen zu überprüfen, kann man den Reifegrad der Umsetzung des Risikomanagements durch eine Reifegradmodell einordnen. Ein solches beschreiben wir näher in Kapitel 7.



6 KI-Risikokompetenz

In unserem Leitfaden „KI-Kompetenzen – Ein praktischer Leitfaden im Sinne der KI-Verordnung der EU“ [6] definieren wir fünf zentrale Kompetenzarten für den verantwortungsvollen Umgang mit KI-Systemen: Fachkompetenz, juristische Kompetenz, ethisch-reflexive Kompetenz, Datenkompetenz und technische KI-Kompetenz. Innerhalb dieser Rahmenstruktur spielt das Bewusstsein für Risiken bereits eine wichtige Rolle. So wird beispielsweise gefordert, dass ethische und juristische Risiken erkannt und bewertet werden können, oder dass Risiken im Umgang mit Daten hinsichtlich Datenschutz, Verzerrungen und Qualität eingeschätzt werden müssen. Diese risikobezogenen Anforderungen innerhalb der allgemeinen KI-Kompetenzarten beziehen sich vor allem auf ein grundlegendes Verständnis von Risiken im jeweiligen Anwendungsbereich. Sie bilden damit eine wichtige Voraussetzung, reichen jedoch nicht aus, um den kontinuierlichen und systematischen Anforderungen an ein Risikomanagement entlang des gesamten Lebenszyklus eines KI-Systems gerecht zu werden.

[6]: Lorenz u. a. (2025), *KI-Kompetenzen – Ein praktischer Leitfaden im Sinne der KI-Verordnung der EU*

Nur auf Basis ausgeprägter Risikokompetenz lässt sich auch Artikel 9 der KI-Verordnung [1] wirksam erfüllen. Dieser fordert unter anderem, dass Anbieter und Betreiber angemessene Maßnahmen zur Beherrschung von Risiken über den gesamten Lebenszyklus eines KI-Systems hinweg sicherstellen. Wie in Kapitel 2 beschrieben, fordern genau das auch etablierte Normen und Empfehlungen, die Risikomanagement als durchgängige Aufgabe über alle Phasen der Entwicklung, Einführung und Nutzung von KI-Systemen hinweg verstehen. Daraus ergibt sich die Notwendigkeit einer differenzierten und strukturierten Risikokompetenz, wie sie im Folgenden beschrieben wird.

[1]: Das Europäische Parlament und der Rat der Europäischen Union (2024), *EU Artificial Intelligence Act*

Risikokompetenz sind somit notwendige Voraussetzung für die Erfüllung der rechtlichen und normativen Vor-

gaben. Die hier eingeführten Risikokompetenz umfasst das Wissen und die Fähigkeiten, Risiken systematisch zu erkennen, zu analysieren und zu bewerten, sowie geeignete Minderungsmaßnahmen zu entwickeln, deren Umsetzung zu begleiten und deren Wirksamkeit kontinuierlich zu überwachen.

Diese erweiterten Fähigkeiten sind insbesondere für Akteure und Rollen erforderlich, die Verantwortung für die Gestaltung, Kontrolle oder Kommunikation des Risikomanagements tragen, wie etwa Risikomanager, Compliance-Verantwortliche oder Qualitätsverantwortliche. Risikokompetenz stellt damit eine spezialisierte Erweiterung der allgemeinen KI-Kompetenzen dar und zielt auf ein methodisch fundiertes, interdisziplinäres Verständnis und die aktive Steuerung von Risiken entlang des gesamten KI-Produktlebenszyklus.

Dieses Kapitel beschreibt einen eigenständigen Kompetenzrahmen für Risikokompetenz und erläutert dessen Nutzen und Anwendung im KI-Kontext inkl. Visualisierung mittels Kompetenzspinnen.

6.1 Niveaus und Rollenbezug

In [6] definieren wir einen praxisorientierten KI-Kompetenzrahmen, mit dessen Hilfe Organisationen systematisch die KI-Kompetenz ihrer Mitarbeiter erfassen, bewerten und dokumentieren können. Dieser Kompetenzrahmen

- benennt zehn an KI-Projekten beteiligte wesentliche Akteure und Rollen,
- identifiziert für verschiedene Kompetenzarten drei aufeinander aufbauende Kompetenzniveaus mit zugehörigen Kenntnissen und Fähigkeiten, und
- ordnet jeder Akteurs-Rolle für jede Kompetenzart ein mindestens notwendiges Kompetenzniveau zu.

Tabelle 6.1 zeigt eine Übersicht über die Akteurs-Rollen aus [6] mit ihren jeweiligen Aufgaben.

Reifegrad KI-Risikomanagement

7

Ein Reifegradmodell für das KI-Risikomanagement beschreibt den Entwicklungsstand einer Organisation im Hinblick auf ihre Fähigkeit, Risiken bei Entwicklung und Einsatz von KI systematisch zu identifizieren, zu bewerten, zu steuern und kontinuierlich zu überwachen. Es zeigt auf, in welchem Maß relevante Strukturen, Prozesse, Kompetenzen und kulturelle Voraussetzungen bereits vorhanden und wirksam sind – und wo gezielter Entwicklungsbedarf besteht.

Das im folgenden vorgestellte Reifegradmodell schafft Transparenz über den aktuellen Reifezustand, unterstützt die Standortbestimmung, ermöglicht die Entwicklung realistischer Zielbilder und dient als Entscheidungsgrundlage für priorisierte Verbesserungsmaßnahmen. Es ist anschlussfähig an zentrale regulatorische und normative Anforderungen, insbesondere an die KI-VO (Art. 9), sowie ISO/IEC 23894, ISO 31000 und NIST AI RMF (vgl. Kapitel 2).

Konzeptionell basiert das Modell auf etablierten wissenschaftlichen Vorgehensweisen zur Reifegradmodellierung von Systemen, insbesondere auf den Arbeiten [7–10]. Diese liefern methodische Leitlinien zur Entwicklung strukturierter Reifegradmodelle entlang definierter Gestaltungsdimensionen und Reifestufen.

Inhaltlich wurde das Modell an die besonderen Anforderungen des KI-Risikomanagements angepasst – insbesondere an die Komplexität KI-spezifischer Risiken, ihre lebenszyklusübergreifende Steuerung sowie Anforderungen an Dokumentation, Nachvollziehbarkeit und Wirksamkeit. Zur Sicherstellung der Praxistauglichkeit wurde das Modell in verschiedenen Organisationen und Workshopformaten erprobt und iterativ weiterentwickelt. Praxisrückmeldungen flossen in die Verfeinerung von Stufenbeschreibungen, Begrifflichkeiten und Anwendungslogik ein.

Nicht-Hochrisiko-KI-Systeme 8

Die KI-VO empfiehlt die Umsetzung von Forderungen für Hochrisiko-Systeme auch andere KI-Systeme[1], um so eine stärkere Verbreitung ethischer und vertrauenswürdiger KI zu fördern. Ein zentraler Baustein hierfür ist die Umsetzung eines Risikomanagementsystems. Das bringt verschiedene Vorteile für Entwicklung und Einsatz eines Nicht-Hochrisiko-Systems mit sich.

1: Erwägungsgründe der KI-VO Satz 165

Akzeptanz und Vertrauen

Auch KI-Systeme, die offiziell Nicht-Hochrisiko-Systeme sind, können vielfältige KI-Risiken und negative ethische Auswirkungen auf Individuen und Gesellschaft haben. Im Sinne der Entwicklung vertrauenswürdiger KI kann es sinnvoll sein, diese von Beginn an systematisch zu betrachten und so Akzeptanz und Vertrauen in das KI-System zu erhöhen.

Unsichere Einstufung

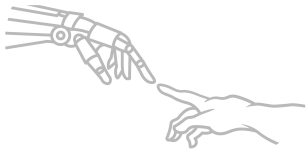
Die Einordnung, ob ein KI-System ein Hochrisiko-KI-System ist oder nicht, kann nicht eindeutig oder unsicher sein und sich im Lauf der Zeit verändern. Zuerst mag beispielsweise eine Einordnung als Nicht-Hochrisiko-KI-System erfolgen, spätere Überprüfungen oder Vorfälle zeigen aber, dass diese Einordnung falsch war oder mittlerweile falsch ist. In solchen Fällen kann es sinnvoll sein, Risiken schon von Beginn an systematisch zu berücksichtigen.

Geringer Zusatzaufwand

Vielleicht sind manche in einer Organisation entwickelte oder eingesetzte KI-Systeme auch Hochrisiko-Systeme, und andere nicht. In solchen Fällen benötigt man für die Hochrisiko-Systeme ein Risikomanagement und die Nicht-Hochrisiko-Systeme lassen sich mit geringem Zusatzaufwand in dieses Risikomanagement integrieren.

Höhere Qualität

Das Risikomanagement ist nach KI-VO Teil des Qualitätsmanagements. Der Hintergrund ist ein enger Zusammenhang zwischen Qualität und Risiken, denn Qualitätsmängel sind auf allen Ebenen Ursache verschiedenster KI-Risiken. Insofern ist Risikomanagement immer

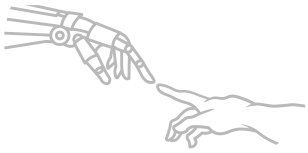


auch Qualitätsmanagement. Viele der genannten Methoden in den verschiedenen Schritten des Risikozyklus beeinflussen die Gesamtqualität des KI-Systems positiv. Dafür sorgen auch die kontinuierliche Überwachung und Aktualisierung.

Mehr Transparenz, Kompetenz und interne Kontrolle

Durch die Einführung von Risikomanagementprozessen und -strukturen auch für Nicht-Hochrisiko-KI-Systeme kann auf verschiedenen weiteren Ebenen ein Mehrwert erzielt werden, wie zum Beispiel: Verbesserte Transparenz und Dokumentation, verbesserte Schulungsmaßnahmen und sachkundigere Nutzung des Systems, erhöhte Akzeptanz des Systems bei Entwicklern, Betroffenen und Nutzern, erhöhtes Vertrauen in das System, Synergieeffekte mit bereits existierenden Risikobetrachtungen in anderen Bereichen, sowie verbesserte Steuerungs- und Weiterentwicklungsmöglichkeiten.

Diese Betrachtungen treffen insbesondere auch auf Allzweck-KI-Systeme zu. Allzweck-KI-Systeme sind nach KI-VO KI-Systeme, die auf einem allgemeinen KI-Modell basieren, das für eine Vielzahl von Zwecken eingesetzt werden kann, sowohl zur direkten Verwendung als auch zur Integration in andere KI-Systeme. Anbieter von Allzweck-KI-Systemen müssen nach der KI-VO verschiedene Dokumentations- und Transparenzpflichten erfüllen und den Urheberschutz gewährleisten. Haben Allzweck-KI-Systeme eine besonders hohe Wirkkraft, Autonomie und Reichweite, so gelten sie als Allzweck-KI-Systeme mit sogenannten systemischen Risiken. Systemische Risiken betreffen vorhersehbare negative Folgen für die öffentliche Gesundheit, die öffentliche Sicherheit oder gesellschaftliche, wirtschaftliche oder politische Strukturen (siehe auch Kapitel 3). Anbieter solcher systemischer Allzweck-KI-Systeme müssen auch Modellbewertungen und Gegentests durchführen und dokumentieren, schwerwiegende Vorfälle verfolgen und melden und Cybersicherheitsschutzmaßnahmen ergreifen. Alle diese Maßnahmen lassen sich auf einfache Weise in den geschilderten Risikomanagementprozess integrieren oder sind dort schon vorhanden. Insbesondere können systemische Risiken vom ersten Schritt des



Risikozyklus an in den Checklisten mitberücksichtigt werden und sind danach automatisch Teil des kompletten Risikozyklus.

Die KI-VO hat zum Ziel, Hochrisiko-Systeme für die Gesamtgesellschaft akzeptabel prüfbar zu gestalten – Maßnahmen zur Überprüfung von Hochrisiko-Systemen sind damit vertrauensbildende Maßnahmen in KI im Allgemeinen. Davon können auch Nicht-Hochrisiko-Systeme profitieren. Zusammengefasst empfehlen wir deshalb auch für Nicht-Hochrisiko-Systeme grundsätzlich die Einführung eines Risikomanagementsystems.

Literaturverzeichnis

[1] Das Europäische Parlament und der Rat der Europäischen Union. *EU Artificial Intelligence Act*. 2024. URL: https://artificialintelligenceact.eu/de/das-gesetz/ (siehe S. 1, 70).

[2] NIST. *Artificial Intelligence Risk Management Framework*. 2024. URL: https://www.nist.gov/itl/ai-risk-management-framework (siehe S. 10).

[3] IEEE. *Ethically Aligned Design*. 2024. URL: https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf (siehe S. 10).

[4] DINMedia. *DIN ISO 31000: 2018-10: Risikomanagement - Leitlinien*. 2018. URL: https://www.dinmedia.de/de/norm/din-iso-31000/294266968 (siehe S. 11).

[5] ISO. *ISO 23894: 2023: Information technology, Artificial intelligence, Guidance on risk management*. 2023. URL: https://www.iso.org/standard/77304.html (siehe S. 11).

[6] Alina Lorenz u. a. *KI-Kompetenzen – Ein praktischer Leitfaden im Sinne der KI-Verordnung der EU*. 2025 (siehe S. 70–72).

[7] Jörg Becker, Ralf Knackstedt und Jens Pöppelbuß. „Entwicklung von Reifegradmodellen für das IT-Management“. In: *Wirtschaftsinformatik* 51.3 (2009), S. 249–260 (siehe S. 81).

[8] Ralf Knackstedt, Jens Poeppelbuss und Jörg Becker. „Vorgehensmodell zur Entwicklung von Reifegradmodellen“. In: *Wirtschaftsinformatik Proceedings*. 2009, S. 535–544 (siehe S. 81).

[9] M.T. Rosemann und T. deBruin. „A model for measuring business process management maturity“. In: *Proceedings of the 13th European Conference on Information Systems Regensburg*. 2005, S. 26–28 (siehe S. 81).

[10] Tonia de Bruin u. a. „Understanding the main phases of developing a maturity assessment model“. In: *ACIS 2005 Proceedings - 16th Australasian Conference on Information Systems*. 2005 (siehe S. 81).

Abbildungsverzeichnis

Tabellenverzeichnis

Universität Augsburg University

Audit- und Wissensplattform für Künstliche Intelligenz