

DIN EN ISO/IEC 23894

ICS 35.240.01

Einsprüche bis 2023-12-06

Entwurf**Informationstechnik –
Künstliche Intelligenz –
Leitlinien für Risikomanagement (ISO/IEC 23894:2023);
Deutsche und Englische Fassung prEN ISO 23894:2023**

Information technology –
Artificial intelligence –
Guidance on risk management (ISO/IEC 23894:2023);
German and English version prEN ISO 23894:2023

Technologies de l'information –
Intelligence artificielle –
Recommandations relatives au management du risque (ISO/IEC 23894:2023);
Version allemande et anglaise prEN ISO 23894:2023

Anwendungswarnvermerk

Dieser Norm-Entwurf mit Erscheinungsdatum 2023-10-06 wird der Öffentlichkeit zur Prüfung und Stellungnahme vorgelegt.

Weil die beabsichtigte Norm von der vorliegenden Fassung abweichen kann, ist die Anwendung dieses Entwurfs besonders zu vereinbaren.

Stellungnahmen werden erbeten

- vorzugsweise online im Norm-Entwurfs-Portal von DIN unter www.din.de/go/entwuerfe bzw. für Norm-Entwürfe der DKE auch im Norm-Entwurfs-Portal der DKE unter www.entwuerfe.normenbibliothek.de, sofern dort wiedergegeben;
- oder als Datei per E-Mail an nia@din.de möglichst in Form einer Tabelle. Die Vorlage dieser Tabelle kann im Internet unter www.din.de/go/stellungnahmen-norm-entwuerfe oder für Stellungnahmen zu Norm-Entwürfen der DKE unter www.dke.de/stellungnahme abgerufen werden;
- oder in Papierform an den DIN-Normenausschuss Informationstechnik und Anwendungen (NIA), 10772 Berlin oder Am DIN-Platz, Burggrafenstr. 6, 10787 Berlin.

Es wird gebeten, mit den Kommentaren zu diesem Entwurf jegliche relevanten Patentrechte, die bekannt sind, mitzuteilen und unterstützende Dokumentationen zur Verfügung zu stellen.

Gesamtumfang 70 Seiten

DIN-Normenausschuss Informationstechnik und Anwendungen (NIA)

Nationales Vorwort

Der Text von ISO/IEC 23894:2023 wurde vom Technischen Komitee ISO/IEC JTC 1 „Information technology“ der Internationalen Organisation für Normung (ISO) erarbeitet und als prEN ISO/IEC 23894:2023 durch das Technische Komitee CEN/CLC/JTC 21 „Künstliche Intelligenz“ übernommen, dessen Sekretariat von DS (Dänemark) gehalten wird.

Das zuständige nationale Normungsgremium ist der Gemeinschaftsarbeitsausschuss NA 043-01-42 GA „DIN/DKE Gemeinschaftsarbeitsausschuss Künstliche Intelligenz“ im DIN-Normenausschuss Informationstechnik und Anwendungen (NIA).

Um Zweifelsfälle in der Übersetzung auszuschließen, ist die englische Originalfassung beigelegt. Die Nutzungsbedingungen für den deutschen Text des Norm-Entwurfes gelten gleichermaßen auch für den englischen Text.

Für die in diesem Dokument zitierten Dokumente wird im Folgenden auf die entsprechenden deutschen Dokumente hingewiesen:

ISO 31000:2018	siehe	DIN ISO 31000:2018-10
ISO 26000:2010	siehe	DIN EN ISO 26000:2021-04
ISO/IEC 29134:2017	siehe	DIN EN ISO/IEC 29134:2020-09

Aktuelle Informationen zu diesem Dokument können über die Internetseiten von DIN (www.din.de) durch eine Suche nach der Dokumentennummer aufgerufen werden.

Nationaler Anhang NA (informativ)

Literaturhinweise

DIN ISO 31000:2018-10, *Risikomanagement — Leitlinien (ISO 31000:2018)*

DIN EN ISO 26000:2021-04, *Leitfaden zur gesellschaftlichen Verantwortung (ISO 26000:2010); Deutsche Fassung EN ISO 26000:2020*

DIN EN ISO/IEC 29134:2020-09, *Informationstechnik — Sicherheitsverfahren — Leitlinien für die Datenschutz-Folgenabschätzung (ISO/IEC 29134:2017); Deutsche Fassung EN ISO/IEC 29134:2020*

– *Entwurf* –

E DIN EN ISO/IEC 23894:2023-11

– Leerseite –

**Informationstechnik – Künstliche Intelligenz – Leitlinien für Risikomanagement
(ISO/IEC 23894:2023)**

Information technology – Artificial intelligence – Guidance on risk management (ISO/IEC 23894:2023)

Technologies de l'information – Intelligence artificielle – Recommandations relatives au management du risque (ISO/IEC 23894:2023)

Inhalt

	Seite
Europäisches Vorwort	4
Vorwort	5
Einleitung	6
1 Anwendungsbereich	7
2 Normative Verweisungen	7
3 Begriffe	7
4 Grundsätze des KI-Risikomanagements	7
5 Rahmenwerk	11
5.1 Allgemeines	11
5.2 Führung und Verpflichtung	11
5.3 Integration	12
5.4 Gestaltung	12
5.4.1 Verstehen der Organisation und ihres Kontextes	12
5.4.2 Artikulieren der Risikomanagementverpflichtung	15
5.4.3 Zuweisung von organisatorischen Rollen, Befugnissen, Verantwortlichkeiten und Rechenschaftspflichten	15
5.4.4 Zuordnung von Ressourcen	15
5.4.5 Einrichten der Kommunikation und Konsultation	15
5.5 Implementierung	15
5.6 Bewertung	16
5.7 Verbesserung	16
5.7.1 Anpassen	16
5.7.2 Fortlaufendes Verbessern	16
6 Risikomanagementprozess	16
6.1 Allgemeines	16
6.2 Kommunikation und Konsultation	16
6.3 Anwendungsbereich, Kontext und Kriterien	16
6.3.1 Allgemeines	16
6.3.2 Festlegen des Anwendungsbereichs	17
6.3.3 Externer und interner Kontext	17
6.3.4 Festlegen von Risikokriterien	18
6.4 Risikobeurteilung	18
6.4.1 Allgemeines	18
6.4.2 Risikoidentifikation	19
6.4.3 Risikoanalyse	21
6.4.4 Risikobewertung	23
6.5 Risikobehandlung	23
6.5.1 Allgemeines	23
6.5.2 Auswahl von Maßnahmen zur Risikobehandlung	23
6.5.3 Erstellen und Implementieren von Plänen zur Risikobehandlung	24
6.6 Überwachen und Überprüfen	24
6.7 Aufzeichnen und Berichten	24
Anhang A (informativ) Ziele	26
A.1 Allgemeines	26
A.2 Verantwortlichkeit	26
A.3 KI-Expertise	26
A.4 Verfügbarkeit und Qualität von Trainings- und Testdaten	26
A.5 Auswirkung auf die Umwelt	27
A.6 Fairness	27
A.7 Instandhaltungsfreundlichkeit	27
A.8 Datenschutz	27

A.9	Robustheit	28
A.10	Sicherheit	28
A.11	Sicherheit	28
A.12	Transparenz und Erklärbarkeit	28
Anhang B (informativ)	Risikoquellen	29
B.1	Allgemeines	29
B.2	Komplexität der Umgebung	29
B.3	Fehlende Transparenz und Erklärbarkeit	29
B.4	Automatisierungsgrad	30
B.5	Risikoquellen in Bezug auf maschinelles Lernen	30
B.6	System-Hardwareprobleme	30
B.7	System-Lebenszyklusprobleme	31
B.8	Technologische Reife	31
Anhang C (informativ)	Risikomanagement und Lebenszyklus von KI-Systemen	32
Literaturhinweise		36

Tabellen

Tabelle 1 — Anwendung von Risikomanagementgrundsätzen auf künstliche Intelligenz	8
Tabelle 2 — Zu prüfende Punkte bei der Festlegung des externen Organisationskontextes	12
Tabelle 3 — Zu prüfende Punkte bei der Festlegung des internen Kontextes einer Organisation	13
Tabelle 4 — Zusätzliche Anleitungen zum Festlegen von Risikokriterien	18
Tabelle C.1 — Risikomanagement und Lebenszyklus von KI-Systemen	32

Europäisches Vorwort

Der Text von ISO/IEC 23894:2023 wurde vom Technischen Komitee ISO/IEC JTC 1 „Information technology“ der Internationalen Organisation für Normung (ISO) erarbeitet und als prEN ISO/IEC 23894:2023 durch das Technische Komitee CEN/CLC/JTC 21 „Künstliche Intelligenz“ übernommen, dessen Sekretariat von DS gehalten wird.

Dieses Dokument ist derzeit zur CEN-Umfrage vorgelegt.

Anerkennungsnotiz

Der Text von ISO/IEC 23894:2023 wurde von CEN als prEN ISO/IEC 23894:2023 ohne irgendeine Abänderung genehmigt.

Vorwort

ISO (die Internationale Organisation für Normung) und IEC (die Internationale Elektrotechnische Kommission) bilden das auf die weltweite Normung spezialisierte System. Nationale Normungsorganisationen, die Mitglieder von ISO oder IEC sind, beteiligen sich an der Entwicklung von Internationalen Normen in Technischen Komitees, die von der jeweiligen Organisation eingerichtet wurden, um spezifische Gebiete technischer Aktivitäten zu behandeln. Auf Gebieten von beiderseitigem Interesse arbeiten die Technischen Komitees von ISO und IEC zusammen. Weitere internationale staatliche und nichtstaatliche Organisationen, die in engem Kontakt mit ISO und IEC stehen, nehmen ebenfalls an der Arbeit teil.

Die Verfahren, die bei der Entwicklung dieses Dokuments angewendet wurden und die für die weitere Pflege vorgesehen sind, werden in den ISO/IEC-Directives, Teil 1 beschrieben. Im Besonderen sollten die für die verschiedenen ISO-Dokumentenarten notwendigen Annahmekriterien beachtet werden. Dieses Dokument wurde in Übereinstimmung mit den Gestaltungsregeln der ISO/IEC-Directives, Teil 2 erarbeitet (siehe www.iso.org/directives oder www.iec.ch/members_experts/refdocs).

Es wird auf die Möglichkeit hingewiesen, dass einige Elemente dieses Dokuments Patentrechte berühren können. ISO und IEC sind nicht dafür verantwortlich, einige oder alle diesbezüglichen Patentrechte zu identifizieren. Details zu allen während der Entwicklung des Dokuments identifizierten Patentrechten finden sich in der Einleitung und/oder in der ISO-Liste der erhaltenen Patenterklärungen (siehe www.iso.org/patents) oder in der IEC-Liste der erhaltenen Patenterklärungen (siehe <http://patents.iec.ch>).

Jeder in diesem Dokument verwendete Handelsname dient nur zur Unterrichtung der Anwender und bedeutet keine Anerkennung.

Für eine Erläuterung des freiwilligen Charakters von Normen, der Bedeutung ISO-spezifischer Begriffe und Ausdrücke in Bezug auf Konformitätsbewertungen sowie Informationen darüber, wie ISO die Grundsätze der Welthandelsorganisation (WTO, en: World Trade Organization) hinsichtlich technischer Handelshemmnisse (TBT, en: Technical Barriers to Trade) berücksichtigt, siehe www.iso.org/iso/foreword.html. In der IEC, siehe www.iec.ch/understanding-standards.

Dieses Dokument wurde vom gemeinsamen Technischen Komitee ISO/IEC JTC 1, *Information technology*, Unterkomitee SC 42, *Artificial intelligence*, erarbeitet.

Rückmeldungen oder Fragen zu diesem Dokument sollten an das jeweilige nationale Normungsinstitut des Anwenders gerichtet werden. Eine vollständige Auflistung dieser Institute ist unter www.iso.org/members.html und www.iec.ch/national-committees zu finden.

Einleitung

Der Zweck des Risikomanagements besteht darin, Werte zu schaffen und zu bewahren. Es verbessert die Leistung, regt Innovationen an und unterstützt das Erreichen von Zielen.

Dieses Dokument ist für die Verwendung in Verbindung mit ISO 31000:2018 vorgesehen. Immer wenn die Leitlinien in diesem Dokument über die Anleitungen in ISO 31000:2018 hinausgehen, wird auf die entsprechenden Abschnitte in ISO 31000:2018 verwiesen, gegebenenfalls gefolgt von einer KI-spezifischen Leitlinie. Um die Zusammenhänge zwischen diesem Dokument und ISO 31000:2018 besser zu verdeutlichen, wird die Abschnittsstruktur von ISO 31000:2018 in diesem Dokument gespiegelt und falls erforderlich durch Unterabschnitte ergänzt.

Dieses Dokument ist in drei Hauptteile untergliedert:

Abschnitt 4: Grundlagen – In diesem Abschnitt werden die zugrundeliegenden Grundsätze des Risikomanagements beschrieben. Der Einsatz von KI erfordert spezifische Erwägungen bezüglich einiger dieser Grundsätze, wie in ISO 31000:2018, Abschnitt 4, beschrieben.

Abschnitt 5: Rahmenwerk – Der Zweck des Risikomanagement-Rahmenwerks ist die Unterstützung der Organisation bei der Integration von Risikomanagement in ihre bedeutende Aktivitäten und Funktionen. Spezifische Aspekte der Entwicklung, der Bereitstellung und des Anbietens sowie der Nutzung von KI-Systemen sind in ISO 31000:2018, Abschnitt 5, beschrieben.

Abschnitt 6: Prozesse – Risikomanagementprozesse umfassen die systematische Anwendung von Managementrichtlinien, -verfahren und -praktiken auf die Tätigkeiten des Kommunizierens, Abstimmens und Festlegens des Kontextes sowie die Bewertung, Behandlung, Überwachung, Überprüfung, Dokumentation und Kommunikation von Risiken. Eine Anpassung dieser Prozesse an KI ist in ISO 31000:2018, Abschnitt 6, beschrieben.

Allgemeine KI-bezogene Ziele und Risikoquellen sind in Anhang A und Anhang B beschrieben. Anhang C enthält ein Beispiel für die Zuordnung zwischen Risikomanagementprozessen und dem Lebenszyklus eines KI-Systems.

1 Anwendungsbereich

Dieses Dokument enthält Anleitungen, wie Organisationen, die Produkte, Systeme und Dienstleistungen entwickeln, einsetzen oder nutzen, bei denen Künstliche Intelligenz (KI) zum Einsatz kommt, mit den spezifischen, mit KI verbundenen Risiken umgehen können. Die Anleitungen haben auch zum Ziel, Organisationen dabei zu unterstützen, das Risikomanagement in ihre KI-bezogene Aktivitäten und Funktionen zu integrieren. Darüber hinaus werden Prozesse für die wirksame Implementierung und Integration des KI-Risikomanagements beschrieben.

Die Anwendung dieser Anleitungen kann für jede Organisation und deren Kontext maßgeschneidert werden.

2 Normative Verweisungen

Die folgenden Dokumente werden im Text in solcher Weise in Bezug genommen, dass einige Teile davon oder ihr gesamter Inhalt Anforderungen des vorliegenden Dokuments darstellen. Bei datierten Verweisungen gilt nur die in Bezug genommene Ausgabe. Bei undatierten Verweisungen gilt die letzte Ausgabe des in Bezug genommenen Dokuments (einschließlich aller Änderungen).

ISO 31000:2018, *Risk management — Guidelines*

ISO Guide 73:2009, *Risk management — Vocabulary*

ISO/IEC 22989:2022, *Information technology — Artificial intelligence — Artificial intelligence concepts and terminology*

3 Begriffe

Für die Anwendung dieses Dokuments gelten die Begriffe nach ISO 31000:2018, ISO/IEC 22989:2022 and ISO Guide 73:2009.

ISO und IEC stellen terminologische Datenbanken für die Verwendung in der Normung unter den folgenden Adressen bereit:

- ISO Online Browsing Platform: verfügbar unter <https://www.iso.org/obp>
- IEC Electropedia: verfügbar unter <https://www.electropedia.org/>

4 Grundsätze des KI-Risikomanagements

Das Risikomanagement sollte auf die Bedürfnisse der Organisation mithilfe einer integrierten, strukturierten und umfassenden Herangehensweise eingehen. Leitgrundsätze ermöglichen es einer Organisation, Prioritäten zu identifizieren und Entscheidungen darüber zu treffen, wie die Auswirkungen von Unsicherheiten auf ihre Ziele zu behandeln sind. Diese Grundsätze gelten für alle Organisationsebenen und Ziele, ob strategisch oder betrieblich.

Systeme und Prozesse setzen üblicherweise eine Kombination verschiedener Technologien und Funktionen für spezifische Anwendungsfälle in unterschiedlichen Umgebungen ein. Das Risikomanagement sollte das Gesamtsystem mit all seinen Technologien und Funktionalitäten und die Auswirkungen auf Umgebung und Stakeholder berücksichtigen.

KI-Systeme können neu auftretende Risiken mit positiven oder negativen Auswirkungen auf die Ziele einer Organisation verursachen oder die Wahrscheinlichkeit des Auftretens vorhandener Risiken ändern. Sie können auch eine spezifische Berücksichtigung durch die Organisation erfordern. In diesem Dokument sind zusätzliche Anleitungen zu Grundsätzen, Rahmenwerk und Prozessen des Risikomanagements beschrieben, die eine Organisation implementieren kann.

ANMERKUNG Die Definitionen des Wortes „Risiko“ unterscheiden sich wesentlich in verschiedenen Internationalen Normen. In ISO 31000:2018 und verbundenen internationalen Normen bezieht sich „Risiko“ auf eine negative oder positive Abweichung von den Zielen. In einigen anderen Internationalen Normen bezieht sich „Risiko“ nur auf potenzielle negative Ergebnisse, beispielsweise auf sicherheitsbezogene Bedenken. Dieser unterschiedliche Fokus kann beim Versuch, einen den Anforderungen entsprechenden Risikomanagementprozess zu verstehen und ordnungsgemäß zu implementieren, häufig für Verwirrung sorgen.

In ISO 31000:2018, Abschnitt 4 sind einige allgemeine Grundsätze des Risikomanagements festgelegt. Zusätzlich zu den Anleitungen in ISO 31000:2018, Abschnitt 4 enthält Tabelle 1 weitere Anleitungen, wie diese Grundsätze bei Bedarf anzuwenden sind.

Tabelle 1 — Anwendung von Risikomanagementgrundsätzen auf künstliche Intelligenz

	Grundsatz	Beschreibung (nach ISO 31000:2018, Abschnitt 4)	Auswirkungen auf die Entwicklung und Nutzung von KI
a)	integriert	Das Risikomanagement ist ein integraler Bestandteil aller Aktivitäten einer Organisation.	Keine spezifischen Anleitungen, die über ISO 31000:2018 hinausgehen.
b)	strukturiert und umfassend	Ein strukturierter und umfassender Risikomanagementansatz trägt zu konsistenten und vergleichbaren Ergebnissen bei.	Keine spezifischen Anleitungen, die über ISO 31000:2018 hinausgehen.
c)	maßgeschneidert	Das Rahmenwerk und die Prozesse des Risikomanagements sind an den externen und internen Kontext einer Organisation angepasst und diesem angemessen sowie mit den Zielen der Organisation verbunden.	Keine spezifischen Anleitungen, die über ISO 31000:2018 hinausgehen.

Tabelle 1 (fortgesetzt)

	Grundsatz	Beschreibung (nach ISO 31000:2018, Abschnitt 4)	Auswirkungen auf die Entwicklung und Nutzung von KI
d)	einbeziehend	Die angemessene und rechtzeitige Beteiligung von Stakeholdern ermöglicht die Berücksichtigung ihrer Kenntnisse, Ansichten und Wahrnehmungen. Dies führt zu verbesserter Erkenntnis und einem fundierten Risikomanagement.	<p>Aufgrund der potenziell weitreichenden Auswirkungen von KI auf Stakeholder ist es wichtig, dass Organisationen in Dialog mit verschiedenen internen und externen Gruppe treten, um sowohl Schäden als auch Vorteile zu kommunizieren und um Rückmeldungen und Erkenntnisse in den Risikomanagementprozess einfließen zu lassen.</p> <p>Organisationen sollten sich auch bewusst sein, dass durch die Anwendung von KI-Systemen zusätzliche Stakeholder involviert werden können.</p> <p>Die Kenntnisse, Ansichten und Wahrnehmungen von Stakeholdern sind unter anderem in folgenden Bereichen von Vorteil:</p> <ul style="list-style-type: none">— Insbesondere das maschinelle Lernen (ML) benötigt für die Erfüllung seiner Ziele häufig geeignete Datensätze. Stakeholder können dazu beitragen, Risiken in Bezug auf die Datenerfassung, die Verarbeitungsvorgänge, die Quelle und Art der Daten und die Nutzung von Daten für bestimmte Situationen oder Ausreißer bei den Datensubjekten zu identifizieren.— Die Komplexität von KI-Technologien verursacht Herausforderungen bezüglich der Transparenz und Erklärbarkeit von KI-Systemen. Die Vielfalt von KI-Technologien verstärkt diese Herausforderungen durch Merkmale wie beispielsweise die unterschiedlichen Arten von Datenmodalitäten, KI-Modell-Topologien und Transparenz- und Berichtsmechanismen, die entsprechend den Bedürfnissen der Stakeholder ausgewählt werden sollten. Stakeholder können dazu beitragen, die Ziele zu identifizieren und die Mittel zur Verbesserung der Transparenz und Erklärbarkeit von KI-Systemen zu beschreiben. In bestimmten Fällen können diese Ziele und Mittel für den Anwendungsfall und die verschiedenen beteiligten Stakeholder verallgemeinert werden. In anderen Fällen kann die Aufteilung der Transparenz-Rahmenwerke und der Berichtsmechanismen durch die Stakeholder je nach Anwendungsfall an relevante Personen (z. B. „Aufsichtsbehörden“, „Geschäftsinhaber“, „Modell-Risikobewerter“) angepasst werden.— Der Einsatz von KI-Systemen für die automatisierte Entscheidungsfindung kann interne und externe Stakeholder direkt betreffen. Diese Stakeholder können beispielsweise ihre Ansichten und Wahrnehmungen dazu äußern, ob möglicherweise eine Überwachung durch Menschen erforderlich ist. Stakeholder können dazu beitragen, Fairnesskriterien zu definieren und zu identifizieren, worin Voreingenommenheit bei der Ausführung von KI-Systemen besteht.

Tabelle 1 (fortgesetzt)

	Grundsatz	Beschreibung (nach ISO 31000:2018, Abschnitt 4)	Auswirkungen auf die Entwicklung und Nutzung von KI
e)	dynamisch	Risiken können auftreten, sich verändern oder verschwinden, wenn sich der externe und interne Kontext einer Organisation verändert. Diese Veränderungen und Ereignisse werden durch das Risikomanagement in angemessener Weise und rechtzeitig vorhergesehen, erkannt, bestätigt und behandelt.	<p>Zur Implementierung der in ISO 31000:2018 bereitgestellten Anleitungen sollten Organisationen organisatorische Strukturen schaffen und Maßnahmen festlegen, um Probleme und Chancen bei aufkommenden Risiken, Trends, Technologien, Einsatzbereichen und Akteuren im Zusammenhang mit KI-Systemen zu identifizieren.</p> <p>Das dynamische Risikomanagement ist aus den folgenden Gründen bei KI-Systemen besonders wichtig:</p> <ul style="list-style-type: none">— KI-Systeme sind durch kontinuierliches Lernen, Verfeinern, Bewerten und Validieren von Natur aus dynamisch. Darüber hinaus haben manche KI-Systeme die Fähigkeit, sich auf Grundlage dieses Kreislaufs anzupassen und zu optimieren und damit selbstständig dynamische Änderungen zu erzeugen.— Die Kundenerwartungen an KI-Systeme sind hoch und können sich potenziell schnell ändern, wenn die Systeme selbst dies tun.— Rechtliche und behördliche Anforderungen im Zusammenhang mit KI ändern sich häufig und werden oft aktualisiert. <p>Auch die Integration in Managementsysteme für Qualität, ökologische Fußabdrücke, Sicherheit, Gesundheitswesen, rechtliche oder unternehmerische Verantwortung oder eine Kombination dieser Themen, die von der Organisation unterhalten werden, kann berücksichtigt werden, um Risiken für die Organisation, Einzelpersonen und Gesellschaften im Zusammenhang mit KI besser zu verstehen und zu behandeln.</p>
f)	beste verfügbare Information	Die Eingaben in das Risikomanagement basieren auf historischen und aktuellen Informationen sowie zukünftigen Erwartungen. Das Risikomanagement berücksichtigt ausdrücklich alle Einschränkungen und Unsicherheiten, die mit solchen Informationen und Erwartungen verbunden sind. Informationen sollten zeitgerecht, verständlich und den relevanten Stakeholdern verfügbar sein.	<p>Berücksichtigt man die Erwartung, dass KI die Art beeinflusst, wie Einzelpersonen mit Technologie interagieren und darauf reagieren, dann ist es für an der Entwicklung von KI-Systemen beteiligten Organisationen ratsam, relevante verfügbare Informationen über den weiteren Einsatz der von ihnen entwickelten KI-Systeme zu verfolgen, während Benutzer von KI-Systemen während der gesamten Lebensdauer dieser Systeme Aufzeichnungen über deren Einsatz führen können.</p> <p>Da es sich bei KI um eine neu aufkommende Technologie handelt, die sich ständig weiterentwickelt, können historische Informationen nur begrenzt verfügbar sein und zukünftige Erwartungen können sich schnell ändern. Organisationen sollten dies berücksichtigen.</p> <p>Wenn überhaupt, sollte die interne Nutzung von KI-Systemen erwogen werden. Die Nachverfolgung der Anwendung von KI-Systemen durch Kunden und externe Benutzer kann durch geistige Eigentumsrechte und vertragliche oder marktspezifische Einschränkungen beschränkt sein. Solche Einschränkungen sollten durch den KI-Risikomanagementprozess erfasst und aktualisiert werden, wenn die Geschäftsbedingungen eine Neubewertung erfordern.</p>

Tabelle 1 (fortgesetzt)

	Grundsatz	Beschreibung (nach ISO 31000:2018, Abschnitt 4)	Auswirkungen auf die Entwicklung und Nutzung von KI
g)	menschliche und kulturelle Faktoren	Menschliches Verhalten und Kultur haben einen wesentlichen Einfluss auf alle Aspekte des Risikomanagements auf allen Ebenen und in jeder Phase.	Organisationen, die an der Gestaltung, Entwicklung oder Auslieferung von KI-Systemen oder einer Kombination davon beteiligt sind, sollten die menschliche und kulturelle Umgebung beobachten, in der sie sich bewegen. Organisationen sollten ihren Fokus darauf richten, zu identifizieren, wie KI-Systeme oder -Komponenten mit vorhandenen gesellschaftlichen Mustern zusammenwirken, wodurch Auswirkungen auf gerechte Ergebnisse, Privatsphäre, Meinungsfreiheit, Fairness, Sicherheit, Beschäftigungsverhältnisse, die Umwelt und die Menschenrechte im Allgemeinen entstehen können.
h)	fortlaufende Verbesserung	Das Risikomanagement wird durch Lernen und Erfahrung fortlaufend verbessert.	Die Identifikation von bisher unbekannten Risiken in Bezug auf den Einsatz von KI-Systemen sollte beim fortlaufenden Verbesserungsprozess berücksichtigt werden. Organisationen, die an der Gestaltung, Entwicklung oder Auslieferung von KI-Systemen oder -Systemkomponenten oder einer Kombination davon beteiligt sind, sollten das KI-Ökosystem in Bezug auf Leistungserfolge, Defizite und gewonnene Erkenntnisse beobachten und ständig neue KI-Forschungsergebnisse und -Techniken (Möglichkeiten zur Verbesserung) verfolgen.

5 Rahmenwerk

5.1 Allgemeines

Der Zweck des Risikomanagement-Rahmenwerks besteht darin, die Organisation dabei zu unterstützen, das Risikomanagement in deren bedeutende Aktivitäten und Funktionen zu integrieren. Es gilt die Anleitung in ISO 31000:2018, 5.1.

Risikomanagement umfasst die Erfassung der für eine Organisation relevanten Informationen, um Entscheidungen zu treffen und auf Risiken einzugehen. Das Steuerungsgremium legt die Gesamtrisikobereitschaft und die organisatorischen Ziele fest, delegiert aber den Entscheidungsfindungsprozess für die Identifizierung, Beurteilung und den Umgang mit Risiken auf Leitungsfunktionen innerhalb der Organisation.

ISO/IEC 38507 [1] beschreibt zusätzliche Erwägungen der Organisationsleitung bezüglich Entwicklung, Kauf oder Einsatz von KI-Systemen. Zu solchen Erwägungen zählen neue Chancen, potenzielle Änderungen der Risikobereitschaft sowie neue Leitungsgrundsätze, um den verantwortungsvollen Einsatz von KI durch die Organisation sicherzustellen. Dies kann zusammen mit den in diesem Dokument beschriebenen Risikomanagementprozessen eingesetzt werden, um die in ISO 31000:2018, 5.2 beschriebene dynamische und sich schrittweise wiederholende organisatorische Integration zu unterstützen.

5.2 Führung und Verpflichtung

Es gilt die Anleitung in ISO 31000:2018, 5.2.

Zusätzlich zu der in ISO 31000:2018, 5.2 bereitgestellten Anleitung gilt Folgendes:

Aufgrund der besonderen Wichtigkeit von Vertrauen und Verantwortlichkeit bei der Entwicklung und Nutzung von KI sollte die oberste Leitung prüfen, auf welche Weise Stakeholder über Grundsätze und Erklärungen zu KI-Risiken und -Risikomanagement informiert werden. Der Nachweis einer entsprechenden Führung und Ver-

pflichtung kann äußerst wichtig sein, um sicherzustellen, dass Stakeholder Vertrauen in die verantwortliche Entwicklung und Nutzung von KI haben.

Die Organisation sollte daher die Veröffentlichung von Erklärungen in Bezug auf ihre Verpflichtung zum KI-Risikomanagement erwägen, um das Vertrauen von Stakeholdern in den Einsatz von KI zu erhöhen.

Die oberste Leitung sollte sich auch darüber im Klaren sein, dass spezialisierte Ressourcen für den Umgang mit KI-Risiken erforderlich sein können, und diese Ressourcen entsprechend bereitstellen.

5.3 Integration

Es gilt die Anleitung in ISO 31000:2018, 5.3.

5.4 Gestaltung

5.4.1 Verstehen der Organisation und ihres Kontextes

Es gilt die Anleitung in ISO 31000:2018, 5.4.1.

Zusätzlich zu der in ISO 31000:2018, 5.4.1 bereitgestellten Anleitung enthält Tabelle 2 weitere Faktoren, die beim Verstehen des externen Organisationskontextes zu berücksichtigen sind.

Tabelle 2 — Zu prüfende Punkte bei der Festlegung des externen Organisationskontextes

Allgemeine, in ISO 31000:2018, 5.4.1 bereitgestellte Anleitung	Zusätzliche Anleitung für Organisationen, die sich mit KI befassen
Organisationen sollten mindestens die folgenden Elemente ihres externen Kontextes prüfen:	Organisationen sollten zusätzlich unter anderem die folgenden Elemente prüfen:
— Soziale, kulturelle, politische, rechtliche, behördliche, finanzielle, technologische, wirtschaftliche und umweltbezogene Faktoren, seien sie internationaler, nationaler, regionaler oder lokaler Art;	— Maßgebliche gesetzliche Belange, einschließlich solcher, die sich spezifisch mit KI befassen. — Leitlinien zur ethischen Nutzung und Gestaltung von KI- und automatisierten Systemen, die von regierungsnahen Gruppen, Aufsichtsbehörden, Normungsorganisationen, der Zivilgesellschaft, Hochschulen und Wirtschaftsverbänden herausgegeben wurden. — KI betreffende, domänenspezifische Leitlinien und Rahmenwerke.
— Wesentliche Schlüsselfaktoren und Trends, welche die Ziele der Organisation beeinflussen;	— Technologische Trends und Fortschritte in den verschiedenen KI-Bereichen. — Gesellschaftliche und politische Auswirkungen beim Einsatz von KI-Systemen, einschließlich sozialwissenschaftliche Leitfäden.
— Beziehungen, Wahrnehmung, Werte, Bedürfnisse und Erwartungen externer Stakeholder;	— Wahrnehmung von Stakeholdern, die durch Probleme wie beispielsweise fehlende Transparenz von KI-Systemen (auch als Undurchsichtigkeit bezeichnet) oder voreingenommene KI-Systeme beeinflusst sein kann. — Erwartungen von Stakeholdern in Bezug auf die Verfügbarkeit spezifischer KI-basierter Lösungen und die Verfahren, wie KI-Modelle verfügbar gemacht werden (z. B. durch Benutzerschnittstellen, Software-Entwicklungskits).

Tabelle 2 (fortgesetzt)

Allgemeine, in ISO 31000:2018, 5.4.1 bereitgestellte Anleitung	Zusätzliche Anleitung für Organisationen, die sich mit KI befassen
<ul style="list-style-type: none">— Vertragliche Beziehungen und Verpflichtungen;	<ul style="list-style-type: none">— Wie der Einsatz von KI und speziell von KI-Systemen, die mit kontinuierlichem Lernen arbeiten, die Fähigkeit der Organisation zum Einhalten vertraglicher Verpflichtungen und Garantien beeinflussen kann. Daraus folgt, dass Organisationen den Geltungsbereich entsprechender Verträge genau prüfen sollten.— Vertragliche Beziehungen während der Gestaltung und Erstellung von KI-Systemen und -Dienstleistungen. Beispielsweise sollten die Eigentumsverhältnisse und Nutzungsrechte von Test- und Trainingsdaten geprüft werden, wenn diese von Dritten bereitgestellt werden.
<ul style="list-style-type: none">— Die Komplexität der Netzwerke und Abhängigkeiten;	<ul style="list-style-type: none">— Die Nutzung von KI kann die Komplexität von Netzwerken und Abhängigkeiten erhöhen.
<ul style="list-style-type: none">— (Anleitungen, die über ISO 31000:2018 hinausgehen).	<ul style="list-style-type: none">— KI-Systeme können bestehende Systeme ersetzen, und in diesem Fall kann eine Beurteilung der Risikovorteile und der Risikoübertragung eines KI-Systems im Vergleich zum bestehenden System erfolgen, wobei sicherheits- und umweltbezogene, technische und finanzielle Probleme in Verbindung mit der Implementierung des KI-Systems geprüft werden.

Zusätzlich zu der in ISO 31000:2018, 5.4.1 bereitgestellten Anleitung enthält Tabelle 3 weitere Faktoren, die für das Verstehen des internen Kontextes einer Organisation zu berücksichtigen sind.

Tabelle 3 — Zu prüfende Punkte bei der Festlegung des internen Kontextes einer Organisation

Allgemeine, in ISO 31000:2018, 5.4.1 bereitgestellte Anleitung	Zusätzliche Anleitung für Organisationen, die sich mit KI befassen
Organisationen sollten mindestens die folgenden Elemente ihres internen Kontextes prüfen:	Organisationen sollten zusätzlich unter anderem die folgenden Elemente prüfen:
<ul style="list-style-type: none">— Vision, Mission und Werte;	<ul style="list-style-type: none">— Keine spezifischen Anleitungen, die über ISO 31000:2018 hinausgehen
<ul style="list-style-type: none">— Leitung, Organisationsstruktur, Rollen und Rechenschaftspflichten;	<ul style="list-style-type: none">— Keine spezifischen Anleitungen, die über ISO 31000:2018 hinausgehen
<ul style="list-style-type: none">— Strategie, Ziele und Grundsätze;	<ul style="list-style-type: none">— Keine spezifischen Anleitungen, die über ISO 31000:2018 hinausgehen
<ul style="list-style-type: none">— Kultur der Organisation;	<ul style="list-style-type: none">— Die Auswirkungen, die ein KI-System auf die Kultur der Organisation haben kann, wenn Verantwortlichkeiten, Rollen und Aufgaben verschoben oder neu eingeführt werden.
<ul style="list-style-type: none">— Von der Organisation übernommene Normen, Leitlinien und Modelle;	<ul style="list-style-type: none">— Alle zusätzlichen internationalen, regionalen, nationalen und lokalen Normen und Leitlinien, die durch den Einsatz von KI-Systemen auferlegt werden.

Tabelle 3 (fortgesetzt)

Allgemeine, in ISO 31000:2018, 5.4.1 bereitgestellte Anleitung	Zusätzliche Anleitung für Organisationen, die sich mit KI befassen
<ul style="list-style-type: none">— Fähigkeiten im Sinne von Ressourcen und Kenntnissen (z. B. Kapital, Zeit, Menschen, geistiges Eigentum, Prozesse, Systeme und Technologien);	<ul style="list-style-type: none">— Die zusätzlichen Risiken für das in der Organisation vorhandene Wissen bezüglich Transparenz und Erklärbarkeit von KI-Systemen.— Durch die Nutzung von KI-Systemen kann sich die Anzahl menschlicher Ressourcen ändern, die für die Umsetzung einer bestimmten Fähigkeit erforderlich sind, oder die Art der erforderlichen Ressourcen kann sich ändern, beispielsweise in Form niedrigerer Qualifikation oder des Verlusts von Expertise, wenn die menschliche Entscheidungsfindung zunehmend durch KI-Systeme gestützt wird.— Die spezifischen Kenntnisse über KI-Technologien und Datenwissenschaft, die für die Entwicklung und Nutzung von KI-Systemen erforderlich sind.— Die Verfügbarkeit von KI-Werkzeugen, -Plattformen und -Bibliotheken kann die Entwicklung von KI-Systemen ermöglichen, ohne dass ein umfassendes Verständnis der Technologie, ihrer Einschränkungen und potenziellen Fallstricke vorhanden ist.— Die Möglichkeit, dass KI zu Problemen und Chancen im Zusammenhang mit dem geistiges Eigentum an spezifischen KI-Systemen führen kann. Organisationen sollten überprüfen, über welches eigene geistige Eigentum sie in diesem Bereich verfügen und wie geistige Eigentumsrechte die Transparenz, Sicherheit und Fähigkeit zur Zusammenarbeit mit Stakeholdern beeinflussen kann, um zu klären, ob und welche Schritte eingeleitet werden sollten.
<ul style="list-style-type: none">— Daten, Informationssysteme und Informationsflüsse;	<ul style="list-style-type: none">— KI-Systeme können zur Automatisierung, Optimierung und Erweiterung der Datenverarbeitung eingesetzt werden.— KI-Systeme nutzen Daten und können daher zusätzliche Anforderungen an die Qualität und Vollständigkeit von Daten und Informationen stellen.
<ul style="list-style-type: none">— Beziehungen zu internen Stakeholdern unter Berücksichtigung ihrer Wahrnehmungen und Werte;	<ul style="list-style-type: none">— Die Wahrnehmung von Stakeholdern, die durch Probleme wie beispielsweise fehlende Transparenz von KI-Systemen oder voreingenommene KI-Systeme beeinflusst sein kann.— Die Bedürfnisse und Erwartungen von Stakeholdern können in höherem Maß durch spezifische KI-Systeme erfüllt werden.— Die Notwendigkeit, Stakeholder über die Fähigkeiten, Fehlermöglichkeiten und den Umgang mit Fehlern von KI-Systemen aufzuklären.

Tabelle 3 (fortgesetzt)

Allgemeine, in ISO 31000:2018, 5.4.1 bereitgestellte Anleitung	Zusätzliche Anleitung für Organisationen, die sich mit KI befassen
— Vertragliche Beziehungen und Verpflichtungen;	<ul style="list-style-type: none">— Die Wahrnehmung von Stakeholdern, die durch unterschiedliche Herausforderungen im Zusammenhang mit KI-Systemen beeinflusst sein kann, wie beispielsweise potenziell fehlende Transparenz und Unfairness.— Die Bedürfnisse und Erwartungen von Stakeholdern können durch spezifische KI-Systeme erfüllt werden.— Die Notwendigkeit, Stakeholder über die Fähigkeiten, Fehlermöglichkeiten und den Umgang mit Fehlern von KI-Systemen aufzuklären.— Die Erwartungen von Stakeholdern zum Schutz personenbezogener Daten und zu grundsätzlichen individuellen und kollektiven Rechten und Freiheiten.
— Gegenseitige Abhängigkeiten und Verbindungen.	— Die Nutzung von KI-Systemen kann die Komplexität von gegenseitigen Abhängigkeiten und Verbindungen erhöhen.

Zusätzlich zur Anleitung in ISO 31000:2018, 5.4.1 sollten Organisationen berücksichtigen, dass für die Nutzung von KI-Systemen eine zusätzliche spezialisierte Ausbildung erforderlich sein kann.

5.4.2 Artikulieren der Risikomanagementverpflichtung

Es gilt die Anleitung in ISO 31000:2018, 5.4.2.

5.4.3 Zuweisung von organisatorischen Rollen, Befugnissen, Verantwortlichkeiten und Rechenschaftspflichten

Es gilt die Anleitung in ISO 31000:2018, 5.4.3.

Zusätzlich zur Anleitung in ISO 31000:2018, 5.4.3 sollten die oberste Leitung und Aufsichtsbehörden, wo angezeigt, die erforderlichen Ressourcen bereitstellen und Personen identifizieren, die:

- die Befugnis zur Behandlung von KI-Risiken haben;
- die Verantwortung für die Einrichtung und Überwachung von Prozessen zur Behandlung von KI-Risiken haben.

5.4.4 Zuordnung von Ressourcen

Es gilt die Anleitung in ISO 31000:2018, 5.4.4.

5.4.5 Einrichten der Kommunikation und Konsultation

Es gilt die Anleitung in ISO 31000:2018, 5.4.5.

5.5 Implementierung

Es gilt die Anleitung in ISO 31000:2018, 5.5.

5.6 Bewertung

Es gilt die Anleitung in ISO 31000:2018, 5.6.

5.7 Verbesserung

5.7.1 Anpassen

Es gilt die Anleitung in ISO 31000:2018, 5.7.1.

5.7.2 Fortlaufendes Verbessern

Es gilt die Anleitung in ISO 31000:2018, 5.7.2.

6 Risikomanagementprozess

6.1 Allgemeines

Es gilt die Anleitung in ISO 31000:2018, 6.1.

Organisationen sollten einen risikobasierten Ansatz zum Identifizieren, Beurteilen und Verstehen von KI-Risiken implementieren, denen sie ausgesetzt sind, und sie sollten entsprechend der Risikohöhe geeignete Behandlungsmaßnahmen ergreifen. Der Erfolg des gesamten KI-Risikomanagementprozesses einer Organisation beruht auf der Identifikation, Festlegung und erfolgreichen Implementierung von eng gefassten Risikomanagementprozessen auf strategischer, betrieblicher, Programm- und Projektebene. Aufgrund von Bedenken, die unter anderem durch die potenzielle Komplexität, fehlende Transparenz und Unvorhersehbarkeit mancher KI-basierter Technologien aufkommen können, sollte den Risikomanagementprozessen auf der KI-System-Projektebene besondere Aufmerksamkeit gewidmet werden. Diese Prozesse auf Systemprojektebene sollten an den Zielen der Organisation ausgerichtet sein, und es sollte ein Informationsaustausch mit anderen Risikomanagementebenen stattfinden. Beispielsweise sollten Eskalationen und gewonnene Erkenntnisse auf KI-Projektebene auch in höhere Ebenen, wie die strategische, betriebliche und Programmebene (und, sofern zutreffend, weitere Ebenen) einfließen.

Anwendungsbereich, Kontext und Kriterien von Risikomanagementprozessen auf Projektebene werden direkt durch die Lebenszyklusphasen des KI-Systems beeinflusst, die im Projektanwendungsrahmen liegen. In Anhang C sind mögliche Zusammenhänge zwischen einem Risikomanagementprozess auf Projektebene und dem Lebenszyklus eines KI-Systems (wie in ISO/IEC 22989:2022 festgelegt) beschrieben.

6.2 Kommunikation und Konsultation

Es gilt die Anleitung in ISO 31000:2018, 6.2.

Die Anzahl der durch ein KI-System betroffenen Stakeholder kann höher sein als anfänglich angenommen, und es können bisher nicht berücksichtigte externe Stakeholder sowie andere Teile einer Gesellschaft betroffen sein.

6.3 Anwendungsbereich, Kontext und Kriterien

6.3.1 Allgemeines

Es gilt die Anleitung in ISO 31000:2018, 6.3.1.

Zusätzlich zur Anleitung in ISO 31000:2018, 6.3.1 sollten in Organisationen, die KI einsetzen, der Anwendungsbereich des KI-Risikomanagements, der Kontext des KI-Risikomanagementprozesses und die Kriterien zur Bewertung der Risikohöhe für die Unterstützung von Entscheidungsfindungsprozessen erweitert werden, um zu ermitteln, wo in der Organisation KI-Systeme entwickelt oder eingesetzt werden. Eine solche Auflistung zur

Entwicklung und Nutzung von KI sollte dokumentiert und in den Risikomanagementprozess der Organisation aufgenommen werden.

6.3.2 Festlegen des Anwendungsbereichs

Es gilt die Anleitung in ISO 31000:2018, 6.3.2.

Der Anwendungsbereich sollte die spezifischen Aufgaben und Verantwortungsbereiche der unterschiedlichen Ebenen einer Organisation berücksichtigen. Darüber hinaus sollten Ziel und Zweck der von der Organisation entwickelten oder genutzten KI-Systeme berücksichtigt werden.

6.3.3 Externer und interner Kontext

Es gilt die Anleitung in ISO 31000:2018, 6.3.3.

Aufgrund der großen potenziellen Auswirkungen von KI-Systemen sollte die Organisation dem Stakeholder-Umfeld bei der Festlegung und Einrichtung des Kontextes eines Risikomanagementprozesses besondere Aufmerksamkeit widmen.

Es sollte erwogen werden, eine Liste von unter anderem den folgenden Stakeholdern zu erstellen:

- die (eigene) Organisation;
- Kunden, Partner und Dritte;
- Lieferanten;
- Endbenutzer;
- Regulierungsbehörden;
- zivile Organisationen;
- Einzelpersonen;
- betroffene Gemeinden;
- Gesellschaften.

Weiterhin sind folgende Punkte beim externen und internen Kontext zu berücksichtigen:

- ob die KI-Systeme Menschen schädigen oder wichtige Dienstleistungen verhindern können (die bei Unterbrechung Leben, Gesundheit oder persönliche Sicherheit gefährden würden), ob sie Menschenrechte verletzen können (z. B. durch unfaire und voreingenommene automatisierte Entscheidungsfindung) oder zu Umweltschäden beitragen können;
- externe und interne Erwartungen bezüglich der gesellschaftlichen Verantwortung einer Organisation;
- externe und interne Erwartungen bezüglich der Umweltverantwortung einer Organisation.

Die Leitlinien nach ISO 26000:2010 [2], in denen Aspekte der sozialen Verantwortung beschrieben sind, sollten als Rahmenwerk für das Verstehen und Behandeln von Risiken Anwendung finden, insbesondere bei Kernthemen wie Organisationsleitung, Menschenrechten, Arbeitsverfahren, Umwelt, fairen Betriebsverfahren, Verbraucherbelangen und der gesellschaftlichen Einbeziehung und Entwicklung.

ANMERKUNG Weitere Hintergrundinformationen zur Vertrauenswürdigkeit sind in ISO/IEC TR 24028:2020 [3] angegeben.