E DIN EN ISO/IEC 23894:2023-11 (D/E)

Erscheinungsdatum: 2023-10-06

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

Inhalt

Seite

Europäisches Vorwort 7

Vorwort 8

Einleitung 9

1 Anwendungsbereich 10

2 Normative Verweisungen 10

3 Begriffe 10

4 Grundsätze des KI-Risikomanagements 10

5 Rahmenwerk 15
5.1 Allgemeines 15
5.2 Führung und Verpflichtung 16
5.3 Integration 16
5.4 Gestaltung 16
5.4.1 Verstehen der Organisation und ihres Kontextes 16
5.4.2 Artikulieren der Risikomanagementverpflichtung 19
5.4.3 Zuweisung von organisatorischen Rollen, Befugnissen, Verantwortlichkeiten und Rechenschaftspflichten 19
5.4.4 Zuordnung von Ressourcen 20
5.4.5 Einrichten der Kommunikation und Konsultation 20
5.5 Implementierung 20
5.6 Bewertung 20
5.7 Verbesserung 20
5.7.1 Anpassen 20
5.7.2 Fortlaufendes Verbessern 20

6 Risikomanagementprozess 20
6.1 Allgemeines 20
6.2 Kommunikation und Konsultation 20
6.3 Anwendungsbereich, Kontext und Kriterien 21
6.3.1 Allgemeines 21
6.3.2 Festlegen des Anwendungsbereichs 21
6.3.3 Externer und interner Kontext 21
6.3.4 Festlegen von Risikokriterien 22
6.4 Risikobeurteilung 23
6.4.1 Allgemeines 23
6.4.2 Risikoidentifikation 23
6.4.3 Risikoanalyse 26
6.4.4 Risikobewertung 28
6.5 Risikobehandlung 28
6.5.1 Allgemeines 28
6.5.2 Auswahl von Maßnahmen zur Risikobehandlung 28
6.5.3 Erstellen und Implementieren von Plänen zur Risikobehandlung 28

6.6 Überwachen und Überprüfen ... 29
6.7 Aufzeichnen und Berichten ... 29

Anhang A (informativ) Ziele ... 30
A.1 Allgemeines ... 30
A.2 Verantwortlichkeit ... 30
A.3 KI-Expertise ... 30
A.4 Verfügbarkeit und Qualität von Trainings- und Testdaten ... 30
A.5 Auswirkung auf die Umwelt ... 31
A.6 Fairness ... 31
A.7 Instandhaltungsfreundlichkeit ... 31
A.8 Datenschutz ... 31
A.9 Robustheit ... 32
A.10 Sicherheit ... 32
A.11 Sicherheit ... 32
A.12 Transparenz und Erklärbarkeit ... 32

Anhang B (informativ) Risikoquellen ... 34
B.1 Allgemeines ... 34
B.2 Komplexität der Umgebung ... 34
B.3 Fehlende Transparenz und Erklärbarkeit ... 34
B.4 Automatisierungsgrad ... 35
B.5 Risikoquellen in Bezug auf maschinelles Lernen ... 35
B.6 System-Hardwareprobleme ... 36
B.7 System-Lebenszyklusprobleme ... 36
B.8 Technologische Reife ... 37

Anhang C (informativ) Risikomanagement und Lebenszyklus von KI-Systemen ... 38

Literaturhinweise ... 42

Tabellen

Tabelle 1 — Anwendung von Risikomanagementgrundsätzen auf künstliche Intelligenz ... 11

Tabelle 2 — Zu prüfende Punkte bei der Festlegung des externen Organisationskontextes ... 16

Tabelle 3 — Zu prüfende Punkte bei der Festlegung des internen Kontextes einer Organisation ... 17

Tabelle 4 — Zusätzliche Anleitungen zum Festlegen von Risikokriterien ... 22

Tabelle C.1 — Risikomanagement und Lebenszyklus von KI-Systemen ... 38