IMPULS RISIKOMANAGEMENT UND -KONTROLLE

KI-Fachkonferenz: „Artificial Intelligence Act – wie Unternehmen sich jetzt einbringen sollten“

Dr. Maximilian Poretschkin, 22. November 2021

Bildquelle: Gorodenkoff/stock.adobe.com

Fraunhofer IAIS

Intelligent Systems that Work!

Fraunhofer IAIS – Künstliche Intelligenz, Machine Learning und Big Data aus Bonn

- Forschung im Paradigma der »hybriden KI« in Partnerschaft mit Exzellenzuniversität Bonn und HBRS
- Nationales Kompetenzzentrum Maschinelles Lernen ML2R, Exzellenzcluster PhenoRob
- Umfangreiches, sofort einsetzbares, nachgewiesen leistungsfähiges Technologie- und IP-Portfolio
- Beratung, 24-7 Umsetzung, Software, Lizensierung, Innovationspartnerschaften, Schulungen
- Kunden und Partner von DAX30 bis Mittelstand
- Netzwerkführung KI.NRW, Fraunhofer-Allianz Big Data und Künstliche Intelligenz, AI4Europe
- Besonderer Fokus auf KI-Absicherung und KI-Zertifizierung

Bildquelle: zapp2photo/stock.adobe.com

Fraunhofer IAIS

Ergebnisse Normungsroadmap KI auf Digital-Gipfel veröffentlicht

Zwei Deliverables zur Etablierung eines Prüfverfahrens:

- Prüfframework, das Vergleichbarkeit von Prüfungen garantiert (und kompatibel mit bestehenden IT-Prüfverfahren ist!)
  - Prozessprüfungen (Standards zur Entwicklung und Betrieb von KI-Systemen)
  - Produktprüfungen (Überprüfung von zugesicherten Eigenschaften)
  - Differenzierte Assurance Levels / Prüftiefen

- Kriterienwerke, welche Anforderungen an Vertrauenswürdigkeit operationalisieren und KI-spezifische Herausforderungen abbilden
  - Use Case Abhängigkeit bei der Formulierung ist Herausforderung (Metriken, Schwellwerte)
  - Völlig neue Prüfwerkzeuge und Prüfmethoden benötigt

Quelle: : DEUTSCHE NORMUNGSROADMAP KÜNSTLICHE INTELLIGENZ (din.de), DIN & DKE, 2019

Standards zu Risikomanagement mit Bezug zu KI

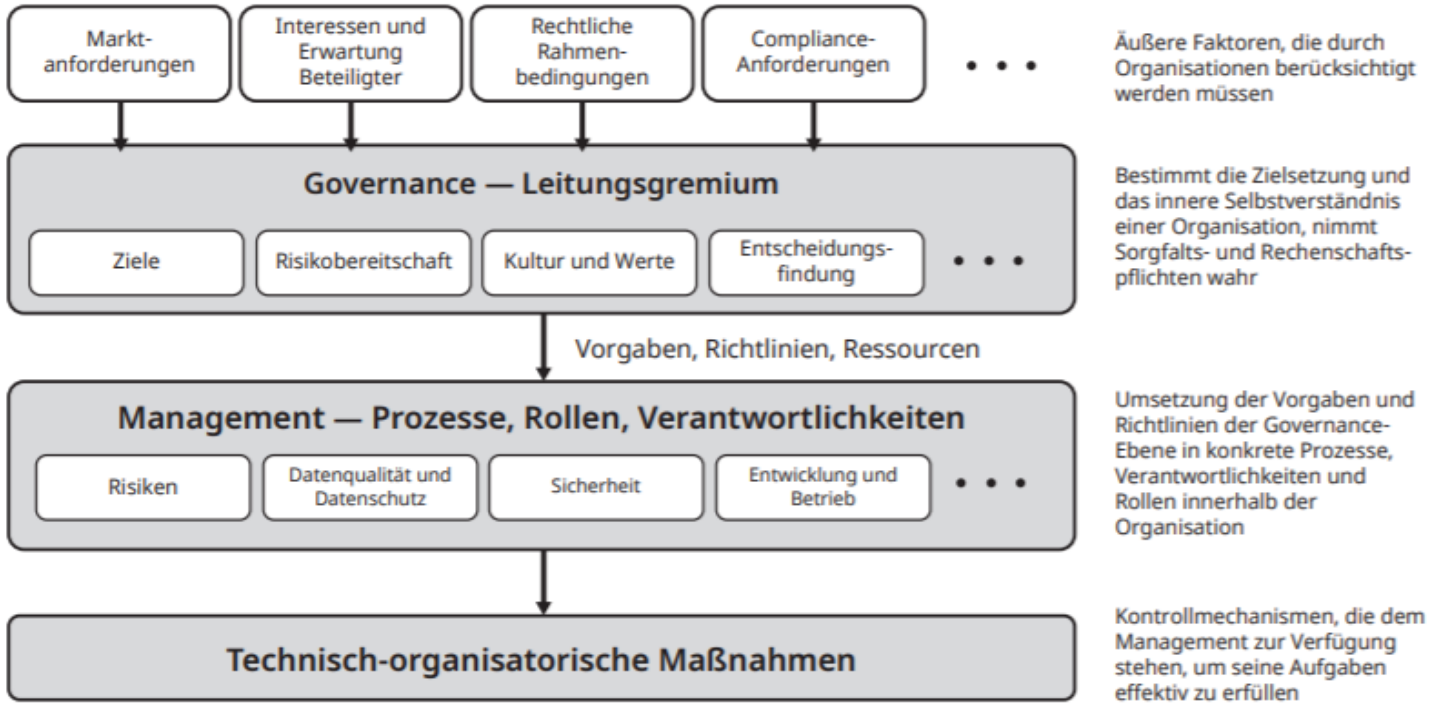
Etablierte Standards werden durch laufende Aktivitäten ergänzt

Dokument	Titel	Kurzbeschreibung
ISO 31000	Risk management – Guidelines	Allgemeine, nicht KI-spezifische Leitlinien zum Risikomanagement. Ansatz für das Behandeln jeglicher Art von Risiko, nicht industrie- oder sektorspezifisch. Basis für ISO/IEC 23894 zum Risikomanagement für KI.
ISO/IEC CD 23894	Information Technology – Artificial Intelligence – Risk Management	Enthält Richtlinien für das Risikomanagement für die Entwicklung und Nutzung von KI-Systemen
ISO/IEC CD 42001	Information Technology — Artificial intelligence — Management system	Allgemeine Beschreibung eines KI-Managementsystems
ISO 14971	Medical devices – Application of risk management to medical devices	Terminologie, Grundsätze und Prozess für das Risikomanagement von Medizinprodukten, eingeschlossen Software als Medizinprodukt

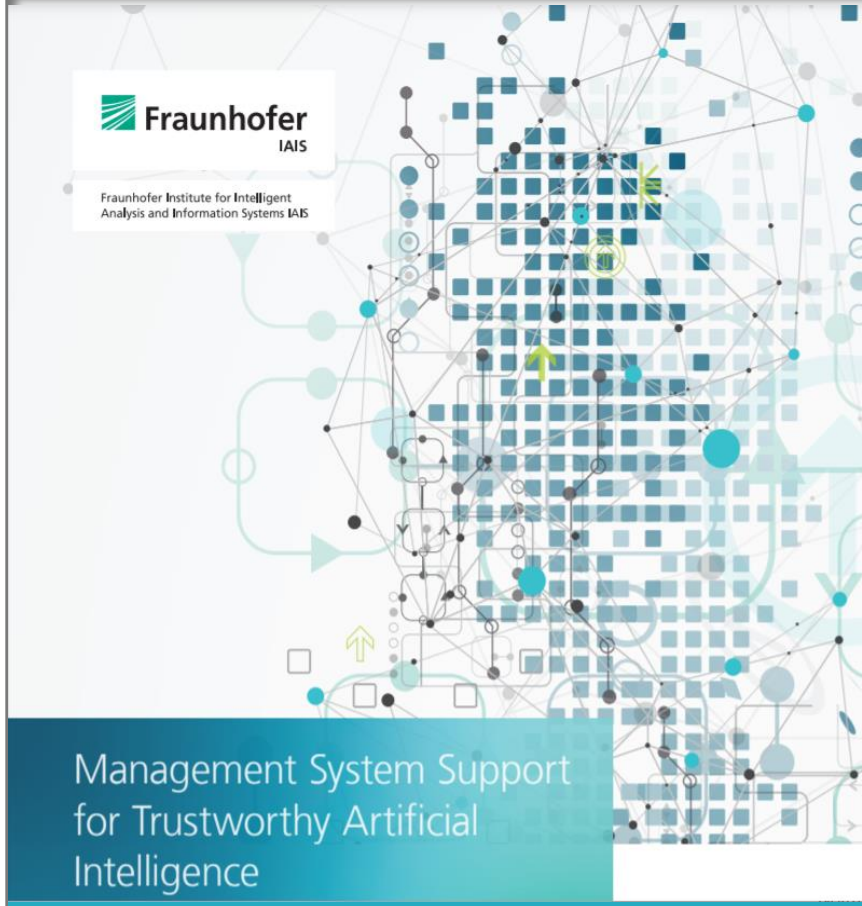
Quelle: : DEUTSCHE NORMUNGSROADMAP KÜNSTLICHE INTELLIGENZ (din.de), DIN & DKE, 2019

KI-Managementsysteme – Vertrauenswürdiger Umgang mit KI durch Organisationen

Überblick über Managementsysteme im Allgemeinen:



Abbildungsquelle: DEUTSCHE NORMUNGSROADMAP KÜNSTLICHE INTELLIGENZ (din.de)



Vergleich des Norm-Entwurf ISO/IEC WD 42001 mit den folgenden KI-Leitplanken:

- Proposal for AI Regulation der Europäischen Kommission
- Assessment List for Trustworthy AI der High-Level Expert Group on AI (HLEG)
- AIC4-Katalog des BSI

KI-Prüfkatalog

- **Schritt 1: Risikoanalyse**
  Umfassende Risikoanalyse entlang der Dimensionen Fairness, Autonomie und Kontrolle, Transparenz, Verlässlichkeit, Sicherheit und Datenschutz
- **Schritt 2: Festlegung von Zielvorgaben**
  Festlegung objektiver, möglichst messbarer Zielkriterien, um die in Schritt 1 identifizierten Risiken abzuschwächen
- **Schritt 3: Auflistung von Maßnahmen**
  Systematische Auflistung von Maßnahmen entlang des Lebenszyklus einer KI-Anwendung, um die in Schritt 2 gesetzten Zielvorgaben zu erreichen
- **Schritt 4: Absicherungsargumentation**
  Erstellung einer stringenten Argumentation, dass die in Schritt 2 formulierten Ziele erreicht wurden

Prüfkatalog ist frei erhältlich unter:
https://www.iais.fraunhofer.de/de/forschung/kuenstliche-intelligenz/ki-pruefkatalog.html

Einsatzbereiche

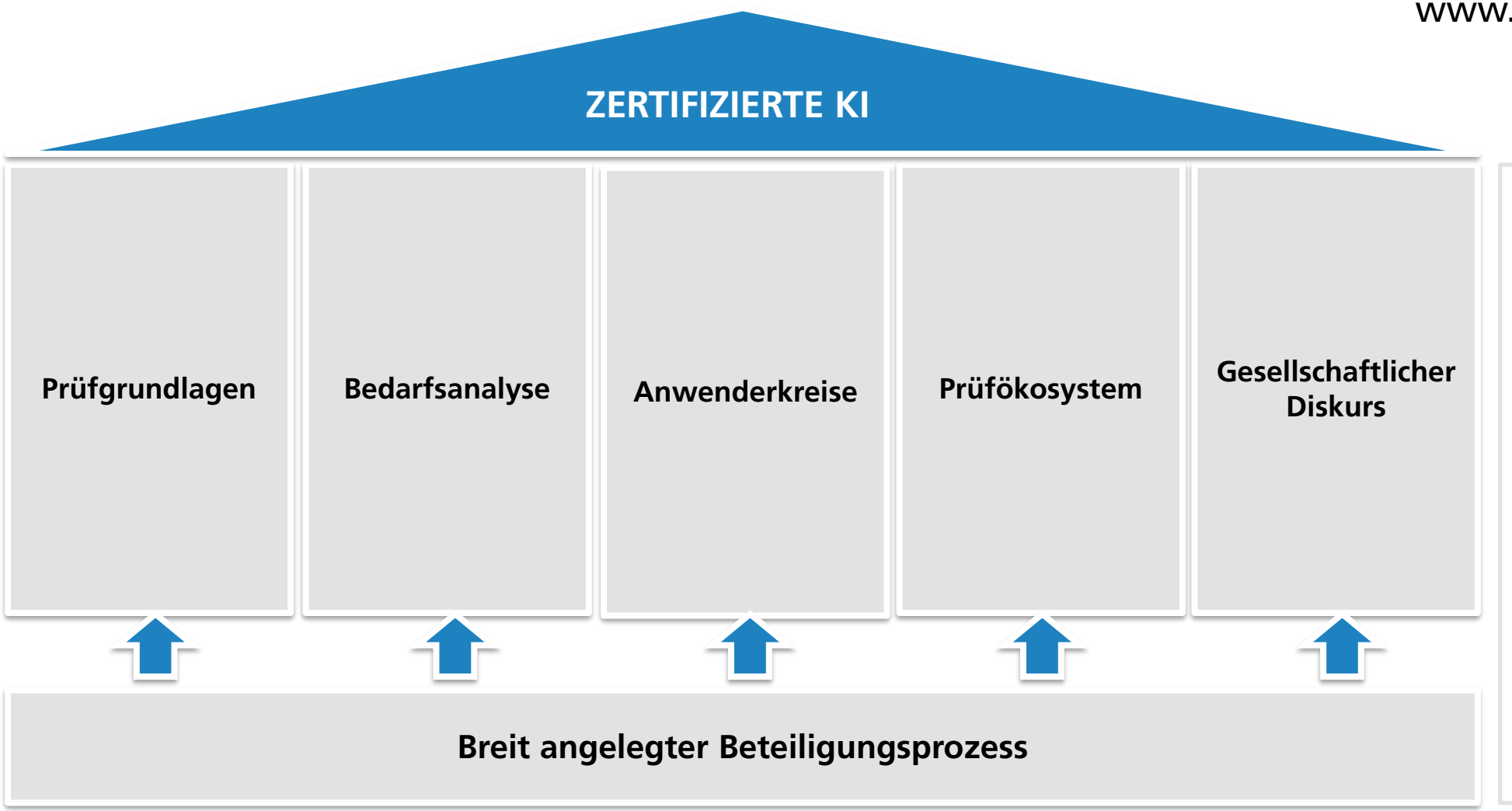
Unser KI-Prüfkatalog unterstützt

- Entwickler*innen bei der Gestaltung und
- KI-Prüfer*innen bei Evaluation und Qualitätssicherung

von KI-Anwendungen.

Struktur Flagship-Projekt ZERTIFIZIERTE KI

ZERTIFIZIERTE KI
Qualität sichern. Fortschritt gestalten.
www.zertifizierte-ki.de



Partner:

Fraunhofer IAIS
Bundesamt für Sicherheit in der Informationstechnik
Universität Bonn
Universität zu Köln
RWTH Aachen University
DIN