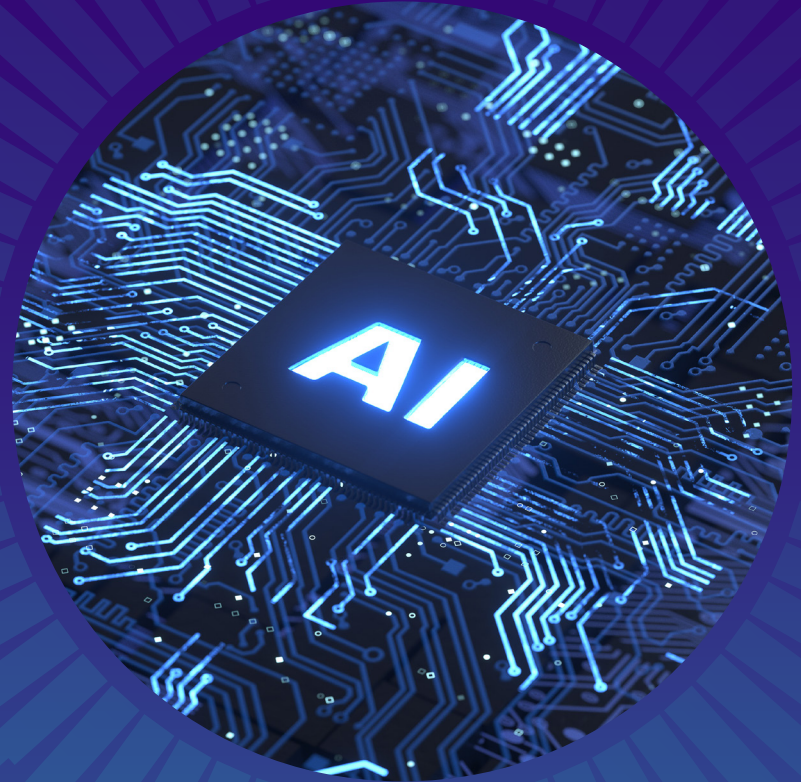




Technical Bulletin

Der Einsatz von Künstlicher Intelligenz (KI) in der
akkreditierten Konformitätsbewertung

Dieses Dokument wurde in Zusammenarbeit erstellt von:



Einleitung

Dieses Joint Technical Bulletin legt gemeinsame Überlegungen von UKAS und DAkkS dar, die eine informative Orientierungshilfe für die verantwortungsvolle Entwicklung, den Einsatz und die Nutzung von auf künstlicher Intelligenz (KI) basierten Technologien und Anwendungen durch Konformitätsbewertungsstellen (KBS) im Rahmen ihrer Geschäftstätigkeit bieten. Es legt keine neuen oder zusätzlichen Anforderungen für die Akkreditierung fest, sondern soll die KBS bei der Identifizierung und Auslegung bestehender geltender Anforderungen unterstützen, die weitgehend technologieneutral formuliert sind. Dieses Dokument ist daher in Verbindung mit den für die Akkreditierung relevanten Normen zu lesen, welche weiterhin die maßgeblichen Dokumente bleiben. Bestehende Anforderungen an KBS bleiben auch im Zusammenhang mit der Nutzung von KI-Systemen von diesem Dokument unberührt.

Informationspflicht gegenüber der nationalen Akkreditierungsstelle bei der Einführung von KI in Konformitätsbewertungsverfahren

Um eine wirksame Kommunikation und die fortlaufende Einhaltung der Akkreditierungsanforderungen zu gewährleisten, wird von der KBS erwartet, ihre nationale Akkreditierungsstelle frühzeitig über wesentliche Änderungen an Ausstattung, Ressourcen oder ihren Konformitätsbewertungsverfahren, -praktiken oder -abläufen zu informieren.

Die Einführung oder Implementierung von KI-Technologien in Konformitätsbewertungsverfahren wird als eine solche wesentliche Änderung betrachtet. Die KBS sind daher verpflichtet, ihre nationale Akkreditierungsstelle zu benachrichtigen, damit potenzielle Auswirkungen überprüft und gegebenenfalls zusätzliche Begutachtungsaktivitäten in Betracht gezogen werden können. Die Mitteilung sollte mindestens Folgendes beinhalten:

- + Eine Kurzbeschreibung des KI-Systems, einschließlich der spezifischen Funktion oder des Verwendungszwecks, für den das System ausgelegt ist, der verwendeten Trainingsdaten/des verwendeten Wissens sowie des Implementierungsansatzes (z. B. Konzepte und Algorithmen des maschinellen Lernens);
- + Informationen darüber, an welcher Stelle des funktionalen Ansatzes bzw. des Konformitätsbewertungsverfahrens das KI-System eingesetzt wird;
- + und den Grad, in dem sich auf das KI-System gestützt wird (z. B. rein administrative Unterstützung, beratende oder entscheidungsunterstützende Funktionen sowie Übertragung von Entscheidungen an die KI).

Da die KBS die Dokumentation auch nutzen kann, um die für ihren KI-Einsatz relevanten Anforderungen zu identifizieren, was eine Voraussetzung für die Sicherstellung der Konformität darstellt, liegt es im eigenen Interesse der KBS, diese Informationen sorgfältig zusammenzustellen.

Berücksichtigung des Kontexts

KI ist eine *General-Purpose-Technologie*, die von KBS auf vielfältige Weise und in unterschiedlichen Kontexten eingesetzt werden kann. Die Ableitung eines einheitlichen oder pauschalen Anforderungskatalogs aus den Level-3-Normen (z. B. ISO/IEC 17025, ISO/IEC 17024, ISO/IEC 17065), der auf alle Anwendungen von KI-Systemen durch KBS anwendbar wäre, ist daher weder praktikabel noch sachgerecht.

Stattdessen ist eine kontextuelle Einordnung unter Berücksichtigung der spezifischen technischen Spezifikationen, des Anwendungsfalls und des Umfelds erforderlich, um zu ermitteln, welche Anforderungen relevant sind und wie diese in der Praxis kontextualisiert und umgesetzt werden müssen. Dies begründet sich darin, dass die spezifische Anwendung eines KI-Systems, dessen technische Merkmale sowie der Kontext, in dem es angewendet wird, für die Anwendbarkeit der Anforderungen aus den jeweiligen Normen entscheidend sind.





Bei der kontextuellen Einordnung von KI in der Konformitätsbewertung müssen mindestens drei Aspekte berücksichtigt werden:

1. Art des KI-Systems

Um relevante Anforderungen zu identifizieren und ihre Anwendbarkeit zu bewerten, muss das KI-System auf technischer Ebene analysiert und beschrieben werden. Zu diesem Zweck sollte international standardisierte Terminologie aus ISO/IEC 22989 *Information technology — Artificial intelligence — Artificial intelligence concepts and terminology*¹ verwendet werden.

Die Analyse muss auf technischer Ebene die vom System ausgeführten Aufgaben, die für Zwecke des Modelltrainings verwendeten und als Eingabe vorgesehenen Daten sowie den Implementierungsansatz umfassen.

+ Anwendungsbereiche:

Die Anwendungsbereiche beziehen sich auf die spezifische Funktion oder den Zweck, für den das KI-System innerhalb der Prozesse der KBS ausgelegt ist. Die Art dieser Aufgaben ist ein wesentlicher Bestimmungsfaktor für potenzielle Risiken und somit auch für die Anwendbarkeit und Priorität bestimmter Anforderungen, die aus den Level-3-Normen abgeleitet werden. So gelten beispielsweise für ein KI-System, das zur automatisierten Dokumentenprüfung eingesetzt wird, andere Anforderungen als für ein System, das für Predictive Maintenance von zu bewertenden Ausrüstungen oder Einrichtungen verwendet wird.

Zur Beschreibung der Aufgaben und Anwendungsbereiche bieten ISO/IEC 22989 Abschnitt 9 (*Fields of AI*) und Abschnitt 10 (*Applications of AI systems*) Beispiele, Konzepte und Terminologie. Da es sich bei KI jedoch um eine *General-Purpose-Technologie* handelt,

deckt die Norm nicht alle möglichen Anwendungsfälle ab, die für KBS relevant sein könnten. Daher ist es in einigen Fällen ratsam, weitere sektor- oder technologiespezifische Normen heranzuziehen.

+ Trainingsdaten/Wissen:

Daten bilden die Grundlage für die Leistung und Zuverlässigkeit von KI-Systemen. Dies betrifft sowohl die zur Entwicklung des Modells verwendeten Trainingsdaten als auch die während des Betriebs verarbeiteten Eingabedaten. Die Art dieser Daten, zum Beispiel ob sie strukturiert (wie bei relationalen Datenbanken) oder unstrukturiert (wie bei Bildern, Audio oder natürlicher Sprache) vorliegen, hat Einfluss auf die technische Komplexität der Datenverarbeitung und wirkt sich auf die Anfälligkeit des Systems für Bias oder Fehler aus. Darüber hinaus können die Herkunft und die Sensibilität der Daten direkt beeinflussen, welche Anforderungen hinsichtlich der Vertraulichkeit oder Informationspflichten gegenüber Kunden Anwendung finden.

+ Implementierungsansatz:

Der Implementierungsansatz bezieht sich auf die spezifischen verwendeten mathematisch-statistischen Methoden und Algorithmen. Dies beinhaltet die Auswahl spezifischer Machine-Learning-Konzepte (z. B. *Supervised Learning* oder *Reinforcement Learning*) und Algorithmen (z. B. neuronale Netze, Entscheidungsbäume oder *Support Vector Machines*). Die dem System zugrunde liegenden Designentscheidungen sind ein wesentlicher Faktor bei der Bestimmung der Anforderungen, insbesondere im Hinblick auf Erklärbarkeit und Validierung. So sind beispielsweise regelbasierte Systeme im Allgemeinen besser erklärbar als Deep-Learning-Modelle, die gegebenenfalls zusätzliche Maßnahmen erfordern, um Transparenzpflichten zu erfüllen. Der Implementierungsansatz bestimmt, ob das System als statisch trainiertes Modell arbeitet oder fortlaufend weiter lernt, wobei Letzteres möglicherweise zusätzliche Sicherheitsmaßnahmen, wie etwa ein strengeres Verfahren der kontinuierlichen Validierung während des Betriebs, erforderlich macht.

ISO/IEC 22989 Abschnitt 5.11 (*Machine learning concepts*) und Abschnitt 5.12 (*Examples of machine learning algorithms*) bieten Beispiele, Konzepte und Terminologie zur Beschreibung des Implementierungsansatzes. Andere Abschnitte der ISO/IEC 22989 enthalten weitere Konzepte und Terminologie zu nicht auf *Machine Learning* basierenden KI-Systemen.

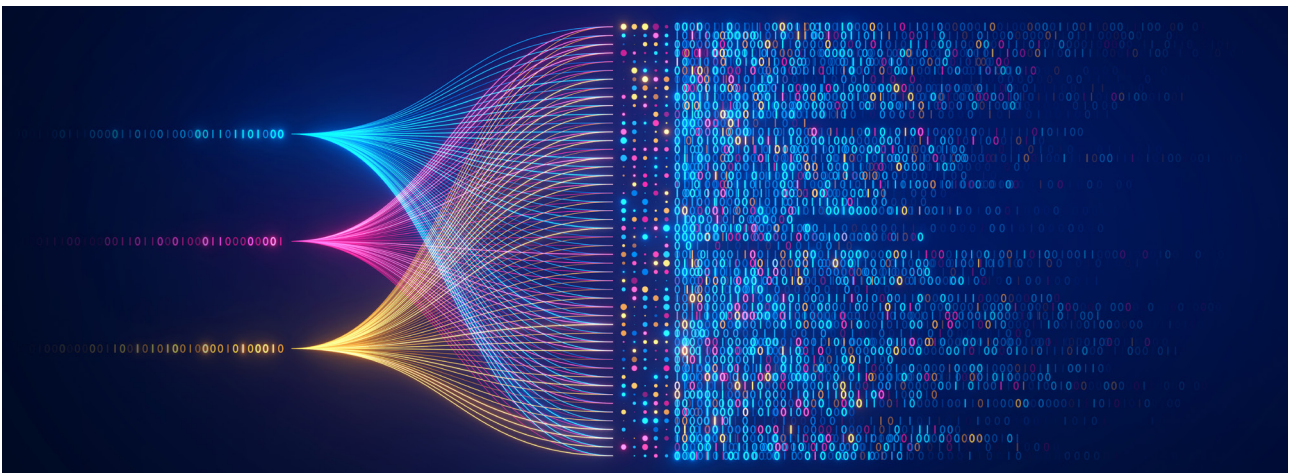
1. Die Norm ist kostenlos unter <https://www.iso.org/standard/74296.html> erhältlich.



2. Der Einsatz des KI-Systems im Konformitätsbewertungsverfahren

Die konkreten Aufgaben, für die das KI-System eingesetzt wird, sind von entscheidender Bedeutung. Systeme, die für administrative Aufgaben (z. B. das Übersetzen von E-Mails) verwendet werden, bergen andere Risiken und unterliegen anderen Anforderungen als solche, die sich direkt auf Entscheidungen über die Konformität auswirken. Die Anwendbarkeit der meisten Anforderungen aus den Level-3-Normen hängt stark davon ab, ob das KI-System in einem der Kernprozessschritte der Konformitätsbewertung eingesetzt wird.

Für die Übertragung administrativer Aufgaben an KI-Systeme, die keinen Einfluss auf Entscheidungen über die Konformität haben, bestehen in der Regel wesentlich geringere Hürden. Es ist unerlässlich, zwischen dem Einsatz von KI in administrativen Funktionen (wie der Übersetzung interner, nichttechnischer Korrespondenz) und einer Anwendung zu unterscheiden, die in den funktionalen Ansatz integriert ist (wie der Übersetzung technischer Dokumentation, die als Eingabe für die Ermittlung dient). Die korrekte Identifizierung, ob ein System einen funktionalen Schritt des Konformitätsbewertungsverfahrens unterstützt oder einem rein administrativen Zweck ohne Auswirkungen auf Konformitätsentscheidungen dient, ist daher eine Voraussetzung für die Bestimmung der relevanten Anforderungen und des zugehörigen Risikoprofils.



3. Grad der Stützung auf das KI-System

Der dritte Faktor, der die Anwendbarkeit der Anforderungen bestimmt, ist der Grad, in dem sich die KBS auf das KI-System stützt. Während die Positionierung des KI-Systems innerhalb des funktionalen Ansatzes oder des Konformitätsbewertungsverfahrens bestimmt, welche Prozessschritte betroffen sind, bestimmt der Grad der Stützung, wie streng die Anforderungen sein müssen. Dieser Faktor hat direkten Einfluss auf das erforderliche Maß an menschlicher Aufsicht und den notwendigen Umfang der Systemvalidierung, um den mit der Nutzung des KI-Systems verbundenen Risiken zu begegnen.

Je stärker sich die KBS bei kritischen Bewertungsschritten auf die Ergebnisse des Systems stützt, desto robuster müssen die erforderlichen Schutzmaßnahmen sein.

Die Rolle des KI-Systems lässt sich in drei Hauptkategorien einteilen:

- + Reine administrative Unterstützung;
- + beratende oder entscheidungsunterstützende Funktionalität;
- + und die Delegation von Entscheidungen an eine KI.

Systeme, die lediglich administrative Unterstützung bieten (die per Definition keinen Einfluss auf Konformitätsentscheidungen haben), bergen das geringste Risiko. Beratende KI-Systeme erfordern ein hohes Maß an Kompetenz der Nutzer sowie eine Validierung der Funktionsweise des KI-Systems. Wenn wesentliche Schritte des Konformitätsbewertungsverfahrens betroffen sind, müssen zudem Prozesse etabliert sein, um Automation Bias und ähnlichen Risiken eines übermäßigen Vertrauens in das System seitens der menschlichen Entscheidungsträger, die diese Unterstützungssysteme nutzen, Rechnung zu tragen. Es ist hervorzuheben, dass endgültige Entscheidungen über die Konformität (die Funktionen Bewertung und Entscheidung, wie in ISO/IEC 17000 definiert) im aktuellen Akkreditierungsrahmen nicht an ein KI-System delegiert werden dürfen.



Vorgehensweise: Wie lassen sich relevante Anforderungen identifizieren und kontextualisieren?

Die Kontextualisierung der relevanten Anforderungen erfordert bei der Anwendung auf KI einen dedizierten Übertragungsschritt. Es ist notwendig, die spezifischen Merkmale und Risiken des betreffenden KI-Systems systematisch mit den Anforderungen der anzuwendenden Level-3-Norm abzugleichen. Dies ermöglicht es der KBS, zu identifizieren, an welchen Stellen die bestehenden Anforderungen der Norm vor dem Hintergrund der spezifisch eingesetzten KI-Technologie ausgelegt werden müssen.

Die Vorgehensweise beginnt mit der Identifizierung der anwendbaren Level-3-Norm und einer Überprüfung des Profils oder der Dokumentation des KI-Systems, wie in den vorangegangenen Abschnitten definiert. Dies umfasst insbesondere dessen technische Merkmale, dessen Rolle im funktionalen Ansatz und den Grad der Stützung auf das System. Die KBS sollte eine Analyse der Level-3-Norm Abschnitt für Abschnitt durchführen und dabei die Systemcharakteristika mit den Anforderungen abgleichen, um deren Anwendbarkeit zu ermitteln.

Diese Analyse zeigt auf, welche generischen Anforderungen durch den jeweiligen Anwendungsfall spezifisch berührt werden. Wenn das System beispielsweise die Weitergabe von Daten an externe Anbieter oder die Nutzung von Kundendaten für das Modelltraining (einschließlich fortlaufendem Lernen im Betrieb) beinhaltet, sind gegebenenfalls die Anforderungen der Norm hinsichtlich der Vertraulichkeit und dem Management von externen Bereitstellungen betroffen. Wird das KI-System in der Ermittlung eingesetzt, muss es als technische Ressource hinsichtlich seiner Funktionstüchtigkeit validiert werden, um sicherzustellen, dass es die Integrität der Konformitätsentscheidung stützt. Dies kann auch eine Aktualisierung der öffentlich zugänglichen Informationen über das Bewertungsverfahren erforderlich machen, um potenzielle Kunden über den Einsatz von KI bei der Ermittlung zu informieren. Darüber hinaus wirkt sich ein solcher Einsatz von KI auf die Anforderungen an das Personal aus. Dies erfordert, dass Kompetenzkriterien und Schulungen erweitert werden, um sicherzustellen, dass das Personal das System effektiv handhaben und Risiken wie eine (systematische) Überbewertung der Verlässlichkeit eines KI-Systems („Automation Bias“) mindern kann.

Die Einhaltung der identifizierten Anforderungen aus den Level-3-Normen kann die Implementierung eigenständiger organisatorischer Prozesse oder spezifischer technischer Schutzmaßnahmen erforderlich machen. Auf prozessualer Ebene könnte dies die Einrichtung wirksamer Mechanismen

zur menschlichen Aufsicht, die verpflichtende Gegenprüfung von KI-generierten Ergebnissen anhand etablierter Verfahren sowie die Sicherstellung umfassen, dass das Personal über die spezifische Kompetenz verfügt, Systemfehler oder Anomalien zu erkennen. Auf technischer Ebene kann die Konformität spezifische Systemkonfigurationen oder Designentscheidungen vorschreiben, wie etwa die lokale Datenverarbeitung zur Gewährleistung der Vertraulichkeit oder die Implementierung von Schutzmaßnahmen, die verhindern, dass Kundendaten zum Training des Modells verwendet werden. Während die Level-3-Normen diese allgemeinen Verpflichtungen festlegen, geben sie keine technischen Spezifikationen vor. Daher sollten die KBS Level-5-Normen heranziehen, um den erwarteten „Stand der Technik“ hinsichtlich technischer Attribute wie Genauigkeit, Transparenz und Robustheit zu bestimmen, sofern Anforderungen an diese Attribute aus der Level-3-Norm abgeleitet werden.

In bestimmten Szenarien finden lediglich einzelne ausgewählte Anforderungen Anwendung. Dies ist beispielsweise der Fall, wenn ein KI-System zur Übersetzung rein administrativer E-Mails eingesetzt wird. In solchen Fällen stellt die KBS (entweder durch Prozesse oder technische Maßnahmen) sicher, dass die Vertraulichkeit gewahrt bleibt und das System nicht außerhalb seines vorgesehenen Anwendungsbereichs verwendet wird, um einen Einsatz für die Übersetzung technischer Dokumentationen, die Entscheidungen über die Konformität beeinflussen könnten, auszuschließen.

Die Analyse des Kontexts kann zudem ergeben, dass eine spezifische Anforderung einen beabsichtigten Anwendungsfall grundlegend untersagt. Wenn beispielsweise eine Level-3-Norm die Verantwortung für eine Zertifizierungsentscheidung explizit einer kompetenten Person zuweist, ist die Ersetzung dieser Person durch ein autonomes, KI-basiertes Entscheidungssystem mit der Norm unvereinbar. In solchen Fällen ergeben sich aus einer Anforderung nicht bloß Rahmenbedingungen für die Implementierung eines Systems, sondern sie fungiert als striktes Ausschlusskriterium für die Delegation von Aufgaben an ein solches System, ungeachtet dessen technischer Ausgestaltung. Folglich müssen KBS solche harten Beschränkungen frühzeitig identifizieren, um sicherzustellen, dass ihr geplanter KI-Einsatz innerhalb der Grenzen liegt, die durch die anzuwendende Akkreditierungsnorm gesetzt sind.





United Kingdom Accreditation Service (UKAS)

Registered Office:

Building Two, Pine Trees, Chertsey Lane,
Staines-upon-Thames, Surrey, TW18 3HR
United Kingdom

Tel: +44 (0)1784 429000

Website: www.ukas.com

Registered in England as a Company
Limited by Guarantee

Company Number: 3076190

Deutsche Akkreditierungsstelle GmbH (DAkkS)

Firmensitz:

Spittelmarkt 10, 10117 Berlin, Germany

Tel: +49 (0)30 670591-0

Email: kontakt@dakks.de

Website: www.dakks.de

Geschäftsführer: Dr.-Ing. Stephan Finke

Vorsitzender Aufsichtsrat:

Bernd Kowalski

Registergericht:

AG Berlin-Charlottenburg

Registernummer: HRB 122846 B

Umsatzsteuer-Identifikationsnummer:

DE815123526

