Bundesministerium
für Digitales und
Staatsmodernisierung

Version 1.0 Oktober 2025

KI-Guide – Ihr Umsetzungsfahrplan für KI-Systeme in der Verwaltung

Vorwort

Beschreibung: Der vorliegende *KI-Guide* dient als **Orientierungshilfe** der Planung und Umsetzung von KI-Projekten in der öffentlichen Verwaltung. Er soll Entscheidenden und Umsetzenden einen Überblick über die wichtigsten Aspekte entlang des KI-Lebenszyklus bieten. Er ist als Wissenskatalog aus **Leitfragen** konzipiert und bietet **Lösungshilfen**, um diese zu beantworten.

Die Lösungshilfen greifen dabei auf Normen und Standards, bewährte Werkzeuge, Erfahrungsberichte und aktuelle Entwicklungen in Technologie und Forschung zurück. Wo sinnvoll, sind sie mit weiterführenden Hinweisen und Links versehen. Der *KI-Guide* soll das **Vertrauen** der Akteure in die erfolgreiche Realisierung Ihrer KI-Vorhaben **stärken** und diese vorantreiben.

Zielgruppen:

Der *KI-Guide* richtet sich vorranging an **Entscheidende** und **Umsetzende**

Entscheidende sind Behördenmitarbeitende, die darüber entscheiden, ob, wie und wofür ein KI-Projekt konzipiert wird. Zu dieser Zielgruppe gehören insbesondere Mitarbeitende in Leitungsfunktionen, beispielsweise Behördenleitungen, Abteilungsleitungen der Z-Abteilung aber auch Mitglieder von KI-Governance-Gremien.

Umsetzende sind Behördenmitarbeitende, die das KI-Projekt umsetzen und dabei auch ethische, technische und rechtliche Grundlagen beachten müssen. Zu dieser Zielgruppe gehören Mitarbeitende der verantwortlichen Organisationseinheiten, insbesondere KI-Projektleitende, IT-Verantwortliche, Entwickelnde und KI-Kompetenzteams.

Jede Frage wurde einer oder mehreren Zielgruppen zugeordnet. Diese Zuordnungen (Tags) sind als Vorschlag zu verstehen. Die tatsächlichen Verantwortlichkeiten in Ihrem Vorhaben können davon abweichen.

Inhalt:

Den Einstieg in den KI-Guide bilden grundlegende Fragen zur Standortbestimmung, die Sie sich zu Beginn stellen sollten. Die einzelnen Aspekte dieser Fragen werden in den folgenden Kapiteln detaillierter erläutert.

Der *KI-Guide* ist entlang der **Phasen des KI-Lebenszyklus** gegliedert:

- **Konzeption und Planung**
- **Datensammlung und Entwicklung**
- **Validierung und Evaluierung**
- **Implementierung und Integration**
- **Wartung und Weiterentwicklung**

Innerhalb der einzelnen Phasen sind die Fragen in **vier Themenschwerpunkte** gruppiert:

Anwendungsfall

Inhalte in diesem Kernbereich umfassen unter anderem die Definition des Ist- und Sollzustands und den Umgang mit relevanten Stakeholdern.

Rechtlicher Rahmen

Inhalte in diesem Kernbereich umfassen unter anderem Aspekte der europäischen Verordnung über künstliche Intelligenz (KI-VO) und der Datenschutz-Grundverordnung (DSGVO).



Technische Umsetzung

Inhalte in diesem Kernbereich umfassen unter anderem die Dokumentation der Datenbasis und die Nutzung von Open-Source-Code.

Verantwortungsvolle Nutzung und Organisation

Inhalte in diesem Kernbereich umfassen unter anderem das Risiko- und Erwartungsmanagement und die Definition eines Notfallplans.

Abhängig von den **konkreten Zielen eines KI-Vorhabens** können bestimmte Fragen mehr oder weniger relevant für die Umsetzung sein.

Sollten Sie weitere Hilfe wünschen, können Sie das BeKI-Beratungsangebot nutzen. Hier geht es zum Beratungsangebot des Beratungszentrums für Künstliche Intelligenz (BeKI) (LINK).

Dieser Leitfaden soll Sie dabei unterstützen, KI-Vorhaben in Ihrer Behörde strukturiert und fundiert anzugehen. Schritt für Schritt erhalten Sie fachliche Hinweise, Methoden und Empfehlungen für ein erfolgreiches Vorgehen.

Gleichzeitig ist uns bewusst: Je nach Ausgangslage, Ressourcen oder Fragestellung kann die Umsetzung im Detail anspruchsvoll sein. Deswegen möchten wir Sie auf das Angebot des Beratungszentrum des Bundes (BZB) im Bundesverwaltungsamt (BVA) aufmerksam machen. Mit 25 Jahren Erfahrung in Projekten, kann das BZB Sie in konkreten Schritten beraten und unterstützen – sei es bei der Prozesserhebung und -gestaltung (2.1.1), bei der nutzendenorientierten Anforderungsaufnahme (2.1.4), der Stakeholderanalyse (2.1.7), bei Personalfragen (2.4.7) oder vielem mehr.

Kontaktieren Sie gerne unmittelbar André Klamrzynski: BZB-Prozesse-Digitales@bva.bund.de. Alle Angebote des BZB finden Sie unter www.beratung.bund.de.

Disclaimer: In der aktuellen Version fokussiert sich der KI-Guide stark auf die Phase „Konzeption und Planung“. Die weiteren Phasen werden angeschnitten, sind aber nicht als vollständig zu betrachten.

1 Standortbestimmung und Vorbereitung

Anhand der folgenden Frage können Sie ermitteln, welcher Startpunkt des KI-Guides sich für Ihren konkreten Fall anbietet.

Wie weit sind Sie bereits in der Konzeption und Planung Ihres KI-Einsatzes?	*Start bei*	
Ich überlege, ob ich KI für einen spezifischen Zweck einsetzen könnte oder ich bin unsicher, ob ich bereits für die konkrete Planung vorbereitet bin.	**1**	**Grundlagen- und Einstiegsfragen**
Ich befinde mich bereits in der konkreten Planung eines KI-Vorhabens.	**2**	**Konzeption und Planung**
Ich habe bereits ein konkretes KI-System für einen festgelegten Einsatzzweck geplant und möchte dieses nun umsetzen.	**3**	**Datensammlung und Entwicklung**



Grundlagen- und Einstiegsfragen

☐ **1.1.1 Ist ein spezifisches Problem oder Ziel definiert, für das KI genutzt werden soll?** Entscheidende

Die Identifizierung des Problems, das durch ein KI-System gelöst werden soll, erfordert eine gründliche Analyse. Umsetzende

Mögliche Methoden sind Interviews mit den von dem Problem Betroffenen oder die Analyse vorhandener Daten zur Erkennung von Mustern, Trends oder Unregelmäßigkeiten, die auf ein Problem hinweisen könnten. Weitere Ansätze sind kreative Brainstorming-Sitzungen mit einem interdisziplinären Team (ggfs. unterstützt durch ein Datenlabor Ihrer Behörde), Use-Case-Analysen, Literaturrecherchen und die Organisation von Workshops mit Expertinnen und Experten.

Zu lösende Probleme können sowohl allgemeine Herausforderungen (z. B. ineffiziente Informationsverarbeitung), als auch spezifische Herausforderungen (z. B. komplizierte Formulare oder unklare Richtlinien in einem bestimmten Verwaltungsprozess) darstellen. KI-Systeme können hier Abhilfe schaffen, indem sie beispielsweise die Datenverarbeitung unterstützen oder Formulare in einfache Sprache übersetzen und Richtlinien automatisch verständlicher formulieren.

☐ **1.1.2 Wurde geprüft, ob sich eine KI sich für die Lösung der Aufgabe eignet und wenn ja, welche?** Entscheidende

Der Einsatz von KI stellt nicht zwangsläufig die effizienteste und passendste Lösung für ein existierendes Problem dar. Die Entscheidung für eine KI-basierte Lösung sollte unter Berücksichtigung verschiedener Kriterien (z. B. Komplexität des Problems, Verfügbarkeit und Qualität von Daten sowie finanzielle Aspekte) getroffen werden. Zudem ist zu prüfen, ob bestehende Technologien oder traditionelle Lösungen effektiv das Problem adressieren können. Umsetzende

Wenn Sie erwägen, eine bestimmte Aufgabe zukünftig mithilfe von KI durchzuführen, sollten Sie sich über die Möglichkeiten und Grenzen der Systeme informieren. Dabei kann die frühzeitige Einbindung von KI-Expertinnen und KI-Experten helfen, sich einen Überblick über die technischen Möglichkeiten zu verschaffen. Zudem empfiehlt sich eine initiale Einschätzung, in welche Risikokategorie gemäß KI-VO das System voraussichtlich fällt.

Diese Information ist relevant, um zu bewerten, ob der Einsatzzweck zulässig ist und ob sich besondere Auflagen für die Einführung ergeben.

☐ **1.1.3 Wurde ein Prozess für die Konzeption des KI-Systems definiert? Wenn ja, an welchem Punkt befindet sich das Vorhaben aktuell?** Entscheidende Umsetzende

Ein typischer Prozess umfasst die folgenden Schritte: Anforderungsformulierung, Konzeptentwicklung, Prototypenerstellung, Validierung, Implementierung und Weiterentwicklung.

☐ **1.1.4 Wurde definiert, wer die Nutzenden und wer die Betroffenen des KI-Systems sind?** Umsetzende

Die Nutzenden eines KI-Systems sind diejenigen, die das System aktiv verwenden, um einen Nutzen daraus zu ziehen oder um bestimmte Ziele zu erreichen. Sie interagieren direkt mit dem System und profitieren von seinen Funktionen oder Dienstleistungen. Betroffene sind hingegen Personen oder Gruppen, die von den Auswirkungen des KI-Systems betroffen sein könnten, auch wenn sie es nicht aktiv nutzen (z. B. indem ihre Daten verarbeitet oder Entscheidungen über ihre Anliegen getroffen werden).

Die Definition der Nutzenden und Betroffenen eines KI-Systems ist ein wichtiger Schritt, um die Bedürfnisse, Anliegen und Interessen der verschiedenen Beteiligten zu verstehen und angemessen darauf reagieren zu können. Potenzielle Risiken und Chancen können so besser zu erkannt und bewertet werden. Zudem können relevante Rechts- und Datenschutzgrundlagen berücksichtigt werden.

☐ **1.1.5 Ist bekannt, welche Daten zur Verfügung stehen?** Entscheidende Umsetzende

Eine frühzeitige Identifikation und Klassifizierung der für das KI-System relevanten Datenquellen ist essenziell. Dazu gehört die Klärung, ob es sich um öffentliche, dienstliche oder um Verschlusssachen, die nur für den Dienstgebrauch bestimmt sind (VS-NfD-Daten), handelt und welche Datenschutzanforderungen (z. B. DSGVO) erfüllt werden müssen. Die Ergebnisse dieser Prüfung haben direkte Auswirkungen auf die Infrastruktur, die Modellentwicklung und die spätere Nutzung des KI-Systems. Chief Data Officer/Scientists in Ihren Behörden können Sie dabei unterstützen. Weitere Informationen zur Datenschutz-Grundverordnung.

☐ **1.1.6 Existieren in meiner Organisation bereits Best Practices zum Einsatz von KI oder gibt es Mitarbeitende, die Erfahrung bei der Konzeption und/oder Umsetzung von KI-Systemen haben?** Umsetzende

Eine Prüfung, ob innerhalb der Organisation bereits Best Practices bestehen oder ob Mitarbeitende mit einschlägiger Erfahrung vorhanden sind, ist eine wichtige Grundlage für die erfolgreiche Umsetzung von KI-Vorhaben. Die Nutzung interner Expertise kann Entwicklungsprozesse beschleunigen, Risiken reduzieren und die Qualität der Ergebnisse verbessern.

Ergänzend dazu sollte im Rahmen des „Marktplatzes der KI Möglichkeiten“ geprüft werden, ob in der Bundesverwaltung bereits nachnutzbare KI-Systeme mit gleichem oder ähnlichem Anwendungsbereich existieren. So kann vorhandenes Wissen ressortübergreifend erschlossen, der Austausch gefördert und die doppelte Entwicklung vergleichbarer Systeme vermieden werden. Dies trägt dazu bei, Ressourcen zu schonen und KI-Systeme schneller und effizienter in Betrieb zu nehmen.

☐ **1.1.7 Sind Mitarbeitende mit IT- und KI-Expertise sowie Mitarbeitende mit Expertise aus dem betroffenen Fachbereich in die Konzeption der Lösung eingebunden?**

Entscheidende

Umsetzende

Um das geplante System sowohl technisch als auch inhaltlich optimal auf die vorliegende Problemstellung abzustimmen, sollten IT- und KI-Expertise sowie die fachliche Expertise aus dem jeweiligen Fachbereich einbezogen werden. Die IT- und KI-Expertise stellt sicher, dass das System mit der bestehenden Infrastruktur technisch umsetzbar und kompatibel ist. Die fachliche Expertise gewährleistet, dass die Lösung die Anforderungen der Nutzenden erfüllt.

Falls intern im betroffenen Fachbereich keine ausreichende IT- und KI-Expertise vorhanden ist, sollte frühzeitig die Unterstützung entsprechender Fachabteilungen oder Anlaufstellen, wie zum Beispiel Datenlabore, in Anspruch genommen werden, um das Vorhandensein der notwendigen Kompetenzen sicherzustellen.

☐ **1.1.8 Ist eine für die Implementierung des KI-Vorhabens geeignete IT-Infrastruktur vorhanden? Falls nicht, gibt es einen Austausch (z. B. über den Marktplatz der KI-Möglichkeiten) über die Anbindung an bestehende Systeme und Infrastrukturen?**

Entscheidende

Umsetzende

Mögliche Infrastrukturen umfassen Server, Grafikprozessoren (GPUs), Dokumentenspeicher und Cloud-Services. Server bilden das Fundament vieler digitaler Systeme, indem sie Daten verarbeiten, speichern und übertragen. GPUs sind besonders effektiv bei rechenintensiven Aufgaben wie maschinellem Lernen und Datenanalyse. Dokumentenspeicher ermöglichen die sichere Aufbewahrung und die schnelle Abrufbarkeit von digitalen Dateien. Die Cloud-Services stellen eine übergeordnete Infrastruktur dar, die eine flexible und skalierbare Umgebung für verschiedene Anwendungen bietet.

Für weiterführende Informationen zu diesen Themen wird der Besuch des „Marktplatzes der KI-Möglichkeiten“ empfohlen. Dort sind eine Vielzahl von Ressourcen, Dienstleistungen und Experten verfügbar.

☐ **1.1.9 Wurden die relevanten Gremien und Interessengruppen innerhalb der Organisation in die Planung und Umsetzung des KI-Systems einbezogen?**

Entscheidende

Umsetzende

Die frühzeitige Einbindung relevanter Gremien und Interessengruppen stellt sicher, dass unterschiedliche Perspektiven berücksichtigt und mögliche Herausforderungen rechtzeitig erkannt werden. Zu diesen Gremien und Gruppen gehören Datenschutzbeauftragte, IT-Fachabteilungen, Personalvertretungen und weitere betroffene Einheiten. Eine transparente Kommunikation und regelmäßige Abstimmung unterstützen eine erfolgreiche Umsetzung und Akzeptanz des KI-Systems.

☐ **1.1.10 Sind die erforderlichen Ressourcen (finanzielle, zeitliche, materielle, personelle) für das Projekt und den (Test-)Betrieb gesichert?**

Entscheidende

Umsetzende

Die Einführung eines KI-Systems erfordert eine sorgfältige Planung sowie die Allokation finanzieller, zeitlicher und personeller Ressourcen. Hierbei ist es wichtig, eindeutige Ziele und Anforderungen festzulegen, realistische Zeitpläne zu entwickeln, das Budget sorgfältig zu planen und die Ressourcen entsprechend zuzuweisen. Auch Zeiten und Aufwände zur Evaluation und Optimierung der Ergebnisse der KI sollten einberechnet werden.

☐ **1.1.11 Wurde bei der Festlegung des Ziels des KI-Systems oder des Prozesses auf Diskriminierungsfreiheit geachtet?** **Umsetzende**

Bei der Einführung eines KI-Systems sollte die Diskriminierung bestimmter Bevölkerungsgruppen unbedingt vermieden werden. Dies ist von Beginn an und durchgehend zu berücksichtigen, zum Beispiel durch die Beantwortung folgender Fragen:

- Ist dieses Ziel für alle Menschen ein gutes Ziel oder nur für einen Teil?
- Wurden die verfügbaren Daten um Diskriminierungen bereinigt beziehungsweise der Umgang hiermit festgelegt?
- Wie kann sichergestellt werden, dass möglichst viele Bevölkerungsgruppen „mitgedacht“ werden?
- Haben etwaige Dienstleister Erfahrungen darin, Diskriminierung in KI zu reduzieren?

2 Konzeption und Planung

Anwendungsfall

☐ **2.1.1 Wurden der Ist- und der Soll-Zustand des durch KI-veränderten Prozesses definiert? Wurden die zu erwartenden Interaktionen zwischen den technischen Komponenten des KI-Systems und den beteiligten sozialen Prozessen beschrieben und aufbereitet?**

Entscheidende

Umsetzende

Die Aufnahme und Visualisierung des Ist-Zustands eines Prozesses hilft dabei, die Wechselwirkungen zwischen Menschen und Technologie aufzuzeigen. Dabei können Schnittstellen und vor allem Schwachstellen des aktuellen Prozesses identifiziert werden, wodurch Ansatzpunkte für die Optimierung des Prozesses offensichtlich werden. Dieses Verständnis ist essenziell, um die Anforderungen an das Soll-System präzise zu formulieren und eine menschenzentrierte Einführung eines KI-Systems zu ermöglichen.

Der Soll-Zustand des Prozesses beschreibt, wie das KI-System genutzt werden soll. Dabei sollten nicht nur der technische Prozess des KI-Systems, sondern auch die menschlichen Schnittstellen im Prozess skizziert werden. Auch im weiteren Prozess sind der Einfluss des KI-Systems auf bestehende Arbeitsabläufe, Interaktionen zwischen Menschen, soziale Dynamiken und menschliches Verhalten zu beachten.

☐ **2.1.2 In welchem Umfang wird erwartet, dass das KI-System zur Steigerung der (betrieblichen) Effizienz beiträgt? Wurden beispielsweise eine Kosten-Nutzenbalance gemessen oder ein Ist- und Soll-Zustand des von KI-veränderten Prozesses definiert?**

Entscheidende

Umsetzende

Die Bestimmung der Effizienzsteigerung durch ein KI-System setzt eine sorgfältige Analyse und Bewertung verschiedener Aspekte der betrieblichen Abläufe und Leistungen voraus. Hierfür müssen potenzielle Effizienz- und Verbesserungspotenziale identifiziert werden (z. B. die Automatisierung manueller Prozesse oder die Optimierung der Ressourcennutzung).

Eine Kosten-Nutzenbalance umfasst die Abwägung der Kosten, die mit der Entwicklung, Implementierung und Nutzung des KI-Systems verbunden sind, im Verhältnis zu den erwarteten Vorteilen und dem Nutzen, den es bringt.

Ein Benchmarking zwischen Ist- und Soll-Zustand ermöglicht ebenfalls eine Gegenüberstellung von Kosten und Nutzen unter Berücksichtigung von Indikatoren wie beispielsweise Zeit, Arbeitspaketen beziehungsweise -schritten und benötigten personellen Ressourcen. Dies hilft außerdem bei der Verbesserung und Weiterentwicklung des Systems.

☐ **2.1.3 Wurde eine Wirtschaftlichkeitsbeurteilung für das KI-System und dessen späteren Betrieb bereits durchgeführt? Beinhaltet diese Wirtschaftlichkeitsbeurteilung eine Zieldefinition, Kostenidentifikation, Nutzenbewertung, einen Zeitrahmen, eine Risikobewertung im Sinne von Investitionsrisiken und eine Berechnung des Return on Investment?**

Entscheidende

Umsetzende

Die Wirtschaftlichkeitsbeurteilung (kurz WiBe) eines KI-Systems im Betrieb ist entscheidend, um sicherzustellen, dass die Investition langfristig rentabel ist und einen Mehrwert für die entsprechende Organisation bietet. Bei der Beschaffung eines KI-Systems sind insbesondere die Kosten für die zugrunde liegende Infrastruktur, die dauerhaften Kosten für Betrieb und Wartung, die Lizenzkosten sowie die Kosten für Dienstleister zu berücksichtigen.

☐ **2.1.4 Sind die Anforderungen an das KI-System bereits festgelegt und dokumentiert?** **Umsetzende**

Ein KI-System muss eine Vielzahl von Anforderungen erfüllen, um effektiv und effizient zu sein und die gewünschte Wirkung zu entfalten. Hierzu zählen Präzision, Skalierbarkeit, Effizienz, Robustheit, Interpretierbarkeit, Datenschutz und Sicherheit sowie Benutzendenfreundlichkeit. Diese Anforderungen sollten quantifiziert und mit den potenziell Nutzenden abgestimmt werden.

Die Bestimmung des erforderlichen Entwicklungsstands und des Autonomiegrads für ein KI-System erfordert eine gründliche Analyse dieser Anforderungen, Ziele und Kontexte, in denen das System eingesetzt wird. Im Sinne der Dokumentation dienen sowohl das „Product Backlog" als auch das Lastenheft dazu, die Anforderungen an ein Softwareprodukt klar zu definieren und zu verwalten.

☐ **2.1.5 Wurde bestimmt, welche Kompetenzen in welchem Umfang für das Projekt benötigt werden? Wenn ja, in welchem Ausmaß sind diese im Team vorhanden?** **Entscheidende**

Für die Einführung eines KI-Systems sind verschiedene Kompetenzen und Fachkenntnisse in den Bereichen Datenwissenschaft, KI/maschinelles Lernen, Softwareentwicklung, Projektmanagement und Produktmanagement im Kernteam des Projektes wichtig. Mitarbeitende im Projekt sollten über ein grundlegendes Verständnis von KI-Technologien, Datenanalysemethoden und Geschäftsprozessen verfügen. **Umsetzende**

Um mit der sich ständig weiterentwickelnden KI-Landschaft Schritt zu halten, ist ein kontinuierliches Lernen unerlässlich. Neben Kompetenzen in der Entwicklung des KI-Systems sind auch Kompetenzen im Bereich Informations- und Change-Management relevant. Weiterführende Informationen finden Sie hier und hier.

☐ **2.1.6 Ist die Beauftragung eines Dienstleisters für den Betrieb notwendig? Wenn ja, sind bereits mögliche Dienstleister bekannt? Sind die Anforderungen definiert, die an einen Rahmenvertragspartner zu stellen sind?** **Entscheidende**

Mögliche Dienstleistungen umfassen Cloud-Services, Hosting und Betrieb, Support, Softwareentwicklung oder die Bereitstellung von Lizenzen und API-Zugängen zu einem KI-Modell. Die Inanspruchnahme solcher Dienstleistungen kann die Art und Menge benötigter Infrastruktur signifikant vereinfachen und mindern. Weitere Informationen finden Sie hier. **Umsetzende**

☐ **2.1.7 Gibt es eine dokumentierte Übersicht über die für das Projekt relevanten Stakeholder und ihre Erwartungen?** **Umsetzende**

Für das Design des Systems ist es wichtig zu wissen, inwieweit die Erwartung der Stakeholder eine ergänzende (Integration neuer Prozesse oder Erweiterung bestehender Prozess) oder eine ersetzende (Ersetzung bestehender menschlicher Arbeitskapazität) Funktion haben. Zudem ist ein Erwartungsmanagement wichtig, das definiert, welche Erwartungen an das KI-System bestehen und das frühzeitig und ehrlich (auch den Endnutzenden) Auskunft darüber gibt, inwieweit diese erfüllbar sind.

☐ **2.1.8 Wurden die Mitarbeitenden des Fachbereichs gezielt für den Umgang mit bzw. den Einsatz von KI sensibilisiert und befähigt?** **Entscheidende**

Die Einführung eines KI-Systems erfordert eine gezielte Sensibilisierung und Befähigung der Mitarbeitenden im Fachbereich. Dies umfasst Schulungen zur Funktionsweise der KI, zu deren Möglichkeiten und Grenzen sowie zu potenziellen Herausforderungen wie Halluzinationen oder Datenschutzrisiken. Ein fundiertes Verständnis stärkt das Vertrauen in den KI-Einsatz und fördert eine sachgerechte Anwendung im Arbeitsalltag. Falls erforderlich, sollten zusätzliche Schulungsangebote oder Leitlinien bereitgestellt werden. **Umsetzende**

☐ **2.1.9 Gibt es für dieses KI-Projekt eine dokumentierte Zuständigkeiten-Struktur in der unter anderem beschrieben ist, welche Akteure für welche Schritte im Konzeptions- und Umsetzungsprozess verantwortlich sind? Berücksichtigt diese Struktur den KI-Lebenszyklus?** **Umsetzende**

Die Entwicklung eines KI-Systems erfordert verschiedene Expertinnen und Experten aus den Bereichen Recht, Technik und Ethik sowie Vertreterinnen und Vertreter der Nutzenden. Diese sind in den jeweiligen Projektphasen in unterschiedlichem Maße erforderlich. Wechselnde, aber zugleich festgelegte Zuständigkeiten je nach Projektphase sind daher wichtig. Es wird empfohlen, dass die zuständigen Personen den gesamten Prozess begleiten.

☐ **2.1.10 Ist beschrieben, wie die KI-Lösung in die bestehenden Arbeitsabläufe und Systeme integriert wird?** **Umsetzende**

Ein KI-System, das nicht in bestehende Strukturen eingegliedert ist, läuft Gefahr, umständlich zu sein (z. B. durch die Notwendigkeit einer mehrmaligen Übertragung von Dokumenten). Eine Integration in bestehende Systeme kann außerdem das Vertrauen der Nutzenden stärken.

☐ **2.1.11 Welche Mechanismen und Kanäle für Feedback sind vorgesehen, um die Nutzendensicht und -erfahrungen kontinuierlich in der Weiterentwicklung des KI-Systems zu berücksichtigen?** **Umsetzende**

Idealerweise sollte Feedback zur Nutzung des KI-Systems dokumentiert werden, damit es an die relevanten Stellen weitergeleitet und für die Optimierung des KI-Systems berücksichtigt werden kann.

☐ **2.1.12 Gibt es einen Plan beziehungsweise bereits implementierte Mechanismen, um das KI-System an einen wachsenden Anwendenden-Kreis oder zunehmend komplexere Aufgabenstellungen anzupassen?** **Umsetzende**

Ein langfristig weiterentwickelbares und effizient nutzbares KI-System muss flexibel genug sein, um sich an sich die Veränderungen von Umständen anzupassen. Zudem ist es für die Bundesverwaltung hilfreich, wenn erfolgreiche Projekte geteilt und nachgenutzt werden können. Besonders wichtig ist, dass das System auch bei großen Datenmengen und bei der Weiterentwicklung von Funktionalitäten nicht an Geschwindigkeit, Nutzendenfreundlichkeit und Qualität einbüßt und ausreichende Zugriffsmöglichkeiten für eine große Anzahl an Nutzenden bietet.

☐ **2.1.13 Wird das KI-System Modell-/Anbieter-agnostisch entwickelt? Ist dokumentiert, wie alternative Modelle/Anbieter im Projektverlauf implementiert werden können?** **Umsetzende**

Aufgrund schneller Entwicklungen im Bereich KI ist eine Anpassungsfähigkeit an neue Systeme von zentraler Bedeutung (sog. Modell-agnostisches System). Dies verringert die Gefahr der Abhängigkeit und erleichtert Lösungen bei Problemen, Fehlern oder nicht zufriedenstellender Performanz.



Rechtlicher Rahmen

☐ **2.2.1 Ist Ihnen die KI-VO bekannt?** **Umsetzende**

Die EU hat mit der KI-VO einen rechtlichen Rahmen für die Entwicklung und Anwendung von KI-Systemen geschaffen. Bitte beachten Sie, dass nicht jedes System, das als KI-System

bezeichnet wird, auch ein KI-System gemäß der KI-VO darstellt. Unter anderem ist etwa erforderlich, dass das KI-System selbstlernend ist.

☐ **2.2.2 Gibt es in Ihrem Team mindestens eine Person, die mit dem rechtlichen Rahmen der Anwendung von KI-Systemen, insbesondere der KI-Verordnung, vertraut ist? Falls nicht, an wen planen Sie, sich mit diesen Fragen zu wenden?** **Umsetzende**

Um die Berücksichtigung des rechtlichen Rahmens sicherzustellen, ist die notwendige Expertise vorzuhalten. Unter den folgenden Links findet sich eine Übersicht rechtlicher Beratungsmöglichkeiten auf deutscher und europäischer Ebene.

☐ **2.2.3 Welche Art der Daten soll das System verarbeiten (öffentliche Daten, dienstliche Daten oder VS-NfD-Daten)?** **Entscheidende**

Die Daten, die ein KI-System verarbeiten soll, sind entscheidend für die Art der Infrastruktur, die genutzt werden kann. VS-NfD-Daten dürfen nur ressortintern gelagert werden. Öffentliche Daten können auch über Cloud-Zugänge verarbeitet werden. **Umsetzende**

☐ **2.2.4 Auf Basis welcher Rechtsgrundlage erfolgt die Nutzung der verfügbaren Daten (z. B. im Hinblick auf Datenschutz, Informationssicherheit, fachliche Vorgaben oder Lizenzen)?** **Entscheidende**

Umsetzende

Hierzu zählen Datenschutzgesetze und Verordnungen wie die Datenschutz-Grundverordnung oder das Bundesdatenschutzgesetz, spezifische Vorschriften für die Informationssicherheit (z. B. BSI-Mindeststandards) sowie die jeweils anwendbaren fachlichen oder lizenzrechtlichen Vorgaben.

☐ **2.2.5 Nutzen Sie im Rahmen der KI-Anwendung personenbezogene Daten?** **Umsetzende**

Sollten Sie personenbezogene Daten zum Training eines KI-Models nutzen, findet die DSGVO Anwendung. Dies bedeutet, dass Sie verschiedene Anforderungen zum Schutz dieser Daten berücksichtigen müssen. Dabei ist zu beachten, dass nicht jedes KI-System personenbezogene Daten verarbeitet. Eine gründliche Prüfung, ob das KI-System Daten verarbeitet, die etwa einen Rückschluss auf die Identität einer Person ermöglichen und somit personenbezogen sind, empfiehlt sich.

Wenn Sie besonders sensible Daten nutzen oder verarbeiten (beispielsweise mit Bezug zu Religion, Gesundheitszustand, politischer Meinung, Herkunft), kann ein erhöhtes Schutzniveau gemäß der DSGVO gelten. Die Einbindung eines / einer Datenschutzbeauftragten oder eines / einer im Datenschutzrecht firmen Person ist daher für die Beachtung der DSGVO, der Vorgaben des Bundesdatenschutzgesetzes sowie von gegebenenfalls Landesdatenschutzgesetzen unerlässlich.

☐ **2.2.6 Sind Ihnen die Leitlinien für den KI-Einsatz in Ihrer Behörde bekannt, beispielsweise die Leitlinien zum KI-Einsatz in der Bundesverwaltung?** **Entscheidende**

Leitlinien unterstützen bei einer zielführenden Umsetzung von KI-Vorhaben, minimieren potenzielle Risiken im Zusammenhang mit dem Einsatz von KI und tragen zur Einhaltung von Vorschriften und Standards bei. **Umsetzende**

Die Leitlinien für den Einsatz Künstlicher Intelligenz in der Bundesverwaltung (KI-Leitlinien) sollen zu einem chancenorientierten und verantwortungsvollen Umgang mit KI-Systemen beitragen. Fünf übergeordnete Leitprinzipien bilden die gemeinsame Wertegrundlage für alle Mitarbeitenden der Bundesverwaltung.

Mit Blick auf die Nutzenden von KI-Systemen sowie die verantwortlichen Organisationseinheiten spezifizieren die Leitsätze die praktischen, ethischen und rechtlichen Rahmenbedingungen für eine verantwortungsvolle Nutzung von KI. Veröffentlicht wurden die Leitlinien im März 2025.

Darüber hinaus können weitere bereichsspezifische Leitlinien nähere Vorgaben zum KI-Einsatz in ihrer Behörde ergänzend festlegen. Es empfiehlt sich, eine organisationsinterne Recherche durchzuführen, falls unklar ist, ob es solche zusätzlichen Leitlinien gibt.

Technische Umsetzung

2.3.1 Basiert der Anwendungsfall auf öffentlich zugänglichen Daten oder sensiblen (ggf. sogar VS-NfD) Daten? Wurden die Implikationen für die Infrastruktur-Auswahl berücksichtigt? Ist bekannt, wann diese Infrastruktur nutzbar zur Verfügung stehen kann?

Entscheidende

Umsetzende

Während öffentliche Daten auch über Cloud-Zugänge verarbeitet werden dürfen, müssen VS-NfD-Daten ressortintern gespeichert sowie Datenverschlüsselungsoptionen und eine Datenschutz-Zertifizierungen bedacht werden. Weitere Informationen finden Sie hier.

Zugriffskontrollen und Berechtigungsmechanismen sowie nach Datensensitivitätsniveau getrennte Datenspeicherorte können sicherstellen, dass bei der Nutzung von KI-Systemen nicht auf geschützte Daten zugegriffen werden kann. Zudem sollte aus der Systemarchitektur ersichtlich sein, auf welchen Datenquellen das System zugreift. Datenzugriffe sind zu protokollieren, um Datenschutzverletzungen sowie Sicherheitslücken zu erkennen.

2.3.2 Ist eine zentrale Nutzung der Anwendung durch mehrere Ressorts möglich?

Entscheidende

Umsetzende

Eine ressortübergreifende Nutzung oder eine Nachnutzung bestehender Systeme und Strukturen sollte zur Schonung von budgetären und zeitlichen Ressourcen in Betracht gezogen werden. Zu prüfende Faktoren sind die Sensibilität von Daten, Infrastruktur, Skalierbarkeit und Zugangsmöglichkeiten. KI-Systeme in der Bundesverwaltung werden zentral im „Marktplatz der KI-Möglichkeiten“ erfasst.

Verantwortungsvolle Nutzung und Organisation

2.4.1 Wurden für die Betroffenen und die Nutzenden geeignete Kommunikationsmaßnahmen geplant oder bereits durchgeführt?

Umsetzende

Für den Erfolg von Projekten ist es entscheidend, die Betroffenen und die Nutzenden mitzunehmen und einzubeziehen. Diese Maßnahmen des Veränderungsmanagements sollten von Beginn an (aufbauend auf einer Stakeholderanalyse) geplant und kontinuierlich weiterentwickelt und umgesetzt werden.

Eine offene und transparente Kommunikation mit allen Zielgruppen und Betroffenen hinsichtlich der Einführung und Nutzung des KI-Systems fördert die Akzeptanz und adressiert frühzeitig mögliche Bedenken. Informationen zur Bereitstellung, Zielsetzung und potenziellen, aus der Nutzung resultieren Änderungen sind von besonderer Bedeutung.

Geeignete Maßnahmen stellen Informationsveranstaltungen, Schulungen, FAQ-Dokumente oder regelmäßige Updates über interne Kommunikationskanäle dar. Besondere Aufmerksamkeit sollte dabei auf die Betroffenen gelegt werden, insbesondere wenn das System Einfluss auf deren Anliegen oder die Verarbeitung ihrer Daten hat.

Darüber hinaus sollte kommuniziert werden, was im Krisenfall (z. B. bei einem Systemausfall) zu tun ist. Weitere Informationen dazu finden Sie hier und hier.

☐ **2.4.2 Gibt es ein Gesamtrisikomanagement für das KI-System?** Entscheidende Umsetzende

Ein Gesamtrisikomanagement betrifft technische, prozessuale, kommunikative, projektbasierte, ethische und rechtliche Aspekte. Risiken sollten identifiziert und mögliche Mitigationsmaßnahmen ergriffen sowie Verantwortlichkeiten für diese benannt werden.

☐ **2.4.3 Wurden KI-spezifische Risiken systematisch erfasst?** Umsetzende

Beispiele für zu erfassende KI-Risiken sind Fairness, Autonomie/Kontrolle, Transparenz, Verlässlichkeit, Sicherheit und Datenschutz. Idealerweise sollten für jeden Aspekt drei Fragen beantwortet werden: Welche Chancen bietet das KI-System? Welche Risiken sind denkbar? Welche Maßnahmen wurden beziehungsweise werden ergriffen, um das Risiko einzudämmen?

Weitere Informationen zu spezifischen Risiken generativer KI finden Sie hier.

☐ **2.4.4 Welchen potenziellen individuellen und gesellschaftlichen Schaden kann das KI-System verursachen? Wie hoch ist die Abhängigkeit von der KI-gestützten Entscheidung und welche Möglichkeiten der Re-Evaluierung gibt es?** Umsetzende

Gefahrenpotenziale existieren vor allem im Hinblick auf eine mögliche Diskriminierung und eine fehlende Erklärbarkeit von Entscheidungen. Wenn Entscheidungen ausschließlich KI-gestützt getroffen werden (z. B. Genehmigungsverfahren), sollte eine menschliche Aufsicht bestehen (human in the loop). Es sollte anhand der Kritikalität der Entscheidung, des Vorhandenseins alternativer Entscheidungsstützen und der Sensitivität (d. h. der Konsistenz von Entscheidungen bei Veränderungen einzelner Parameter und Informationen) evaluiert werden.

Bei weniger folgenreichen Anwendungsfällen kann ein Mindestmaß an Kontrolle darin bestehen, den Einsatz des KI-Systems jederzeit beenden zu können. Hierzu sollte geprüft werden, inwiefern diese Möglichkeit sichergestellt ist (siehe dazu 4.2.1.).

☐ **2.4.5 Wurde definiert, bei wem die Entscheidungsverantwortung liegt? Sind Kontrollpunkte (menschlich sowie technisch) vorgesehen?** Entscheidende Umsetzende

Die Verantwortung kann bei Entwicklerinnen und Entwicklern, Betreiberinnen und Betreibern oder Entscheidungsträgerinnen und Entscheidungsträgern liegen und sollte klar definiert sein. Eine regelmäßige Validierung der Ergebnisse empfiehlt sich. Nutzende sollten auf ihre Kontroll- und Entscheidungsverantwortung hingewiesen sowie bei Bedarf entsprechend geschult werden.

☐ **2.4.6 Wurden Risiken, die aus der Kombination mit anderen IT-Anwendungen sowie dem Nutzendenverhalten resultieren können, berücksichtigt?**

Umsetzende

Das KI-System sollte nicht durch Nutzende manipulierbar sein. Dies betrifft sowohl versehentliche (z. B. durch Formulierung des Prompts) als auch willentliche Manipulation. Zudem ist darauf zu achten, dass die Integration in bestehende Systeme deren Datenschutz und Sicherheit nicht gefährdet.

☐ **2.4.7 Kann die Nutzung von KI-Systemen Auswirkungen auf die betroffenen Arbeitsplätze haben?**

Entscheidende

Umsetzende

Es ist zu überlegen, ob und, wenn ja, welche Aufgaben durch die KI-Nutzung wegfallen, sich verändern oder neu hinzukommen. Je nach Grad der Veränderung des Arbeitsplatzes kann dies Auswirkungen auf die tarifliche Eingruppierung beziehungsweise Dienstpostenbewertung haben. Die entsprechenden Konsequenzen sollten frühzeitig mitgedacht werden.

Sollten die Aufgaben durch das KI-System weniger Personal erfordern, so ist der neue Personalbedarf zu ermitteln. Weitere Informationen dazu finden Sie hier.

3 Datensammlung und Entwicklung

Rechtlicher Rahmen

☐ **3.1.1 Haben Sie die erforderlichen Anforderungen der KI-VO für das betreffende KI-System beachtet?** **Umsetzende**

Die KI-VO beinhaltet insbesondere für Anbieter und Betreiber wichtige Pflichten. Da die KI-VO einen risikobasierten Ansatz verfolgt, müssen Sie zunächst Ihr System einer der Kategorien zuordnen (sog. Risikoeinschätzung). Welche Pflichten beachtet werden müssen, richtet sich nach der entsprechenden Risikoeinstufung. Beachten Sie, dass manche KI-Systeme verboten sind oder als solche mit hohem Risiko zu qualifizieren sind. Weitere Informationen finden Sie auf der Seite der Bundesnetzagentur.

☐ **3.1.2 Wie wird der Schutz (besonders) personenbezogener Informationen und die Einhaltung von Datenschutzbestimmungen im Rahmen des Datenmanagements und der KI-Verarbeitungsprozesse gewährleistet? Welche fachlichen Vorschriften müssen über den Datenschutz hinaus eingehalten werden?** **Entscheidende** **Umsetzende**

Um den Schutz personenbezogener (insbesondere sensibler) Daten zu gewährleisten, sollte sichergestellt werden, dass das System ausschließlich auf die für die jeweilige Verarbeitung notwendigen Daten zugreift. Jede Verarbeitung bedarf einer klaren Rechtsgrundlage. Datenschutzrechtliche Prinzipien wie „Privacy by Design“ sind frühzeitig in die Systemarchitektur und Datenprozesse zu integrieren. Technische und organisatorische Maßnahmen wie Zugriffskontrollen, Anonymisierung, Pseudonymisierung und Verschlüsselung sensibler Daten tragen ebenfalls maßgeblich zur Einhaltung der Datenschutzanforderungen bei. Darüber hinaus erleichtert die Trennung von öffentlichen und geschützten Daten die Zugriffssteuerung und reduziert das Risiko von Datenlecks.

Die Möglichkeit einer **systematischen Anonymisierung** von Daten vor ihrer Nutzung sollte geprüft und – wenn möglich – umgesetzt werden. Anonymisierte Daten unterliegen geringeren rechtlichen Auflagen und können die Verarbeitung und Analyse erheblich vereinfachen, insbesondere im Rahmen von KI-Anwendungen.

Zur Analyse potenzieller Risiken und Auswirkungen auf die Rechte und Freiheiten betroffener Personen sollte eine **Datenschutzfolgeabschätzung** durchgeführt werden, insbesondere dann, wenn die Datenverarbeitung voraussichtlich ein hohes Risiko birgt. Sie dient der Identifikation und Bewertung datenschutzrelevanter Gefahren sowie der Ableitung geeigneter Schutzmaßnahmen.

Beim **Training von KI-Systemen** ist sicherzustellen, dass ausschließlich zulässige Daten verwendet werden – also keine Daten, deren Nutzung gegen Eigentums-, Urheber- oder Datenschutzrechte verstößt. Die Datenbasis muss transparent und nachvollziehbar sein. Eine unkontrollierte Nutzung externer oder nicht validierter Datenquellen (z. B. aus dem offenen Internet) ist zu vermeiden. Stattdessen sollte sorgfältig evaluiert werden, welche Datensätze inhaltlich und rechtlich geeignet sind, um ein belastbares Trainingsdatenset zu erstellen.

Die genannten Maßnahmen orientieren sich an zentralen Vorgaben der DSGVO und dienen exemplarisch der Verdeutlichung datenschutzkonformer Umsetzungsschritte. Sie stellen jedoch keine abschließende Aufzählung aller rechtlichen Anforderungen dar – je nach Art und Umfang der Datenverarbeitung können zusätzliche gesetzliche Vorgaben und spezifische Prüfpunkte relevant sein, die ebenfalls zu berücksichtigen sind.

☐ **3.1.3 Wurden Maßnahmen implementiert, um die Informationssicherheit und die Integrität der Daten innerhalb des KI-Systems entlang der Vorgaben unter anderem des Bundesamt für Sicherheit in der Informationstechnik (BSI) zu gewährleisten?** Entscheidende Umsetzende

Relevant sind Verschlüsselungen, Zugriffskontrollen, Authentifizierung, Protokollierung und Überwachung.

Technische Umsetzung

☐ **3.2.1 Ist die verwendete Datenbasis ausreichend dokumentiert?** Umsetzende

Eine ausreichende Dokumentation umfasst Angaben zur Herkunft, Struktur, Beschreibung, Aktualisierungshistorie, Datenverarbeitung, zu rechtlichen und ethischen Überlegungen, zur Qualitätssicherung sowie bekannten Limitationen oder Fehler. Ein umfassendes Verständnis der verwendeten Datenbasis ist notwendig, um die Ergebnisse von KI-Systemen nachvollziehen zu können. Auch unterstützt eine solche Dokumentation eine eventuelle Fehlersuche (z. B. beim Auftreten von Bias).

☐ **3.2.2 Wie wird die Datenbereitstellung gewährleistet?** Umsetzende

Oft sind große Datenmengen für das Training einer Anwendung nötig. Sollten keine internen Daten in ausreichendem Umfang zur Verfügung stehen, können die Datenlabore z.B. zur Erzeugung von Trainingsdaten oder zur Aufbereitung vorhandener Daten unterstützen. Darüber wird eine Prüfung empfohlen, wie öffentlich zur Verfügung gestellte Datensätze (Open Data) von Portalen wie GovData genutzt werden können.

☐ **3.2.3 Wird der Prozess der technischen Umsetzung (inklusive notwendiger Implementierungsentscheidungen) transparent dokumentiert?** Umsetzende

Entscheidungen hinsichtlich der Auswahl der Technologie, Plattformen, Modelle, Algorithmen, Datenquellen und Zugangsbeschränkungen zu dokumentieren und zu begründen sind. Für Use Cases, die unter die KI-VO fallen, wird an dieser Stelle spezifiziert, welche Inhalte die technische Dokumentation umfassen muss.

☐ **3.2.4 Arbeitet das Projekt mit Open-Source-Code?** Umsetzende

Die Arbeit mit Open-Source-Code fördert Innovation, Sicherheit, Transparenz und Langlebigkeit, da die Entwicklung des KI-Systems auch für Außenstehende einsehbar und somit zu verstehen ist. Auch besteht keine Abhängigkeit von bestimmten Anbietern und es fallen keine Lizenzgebühren an. Jedoch sind auch bei Open-Source-Code die Rechte und Pflichten aus den entsprechenden Lizenzvereinbarungen zu beachten.

☐ **3.2.5 Wie werden Datenqualität und -management (auch im weiteren Projektverlauf) sichergestellt?** Umsetzende

Durch Datenstandardisierung, -validierung und -bereinigung werden Duplikate und Inkonsistenzen vermieden. Die Automatisierung dieser Prozesse minimiert die Möglichkeit menschlicher Fehler. Zudem stellen eine regelmäßige Aktualisierung der Datenquellen und Trainingsdaten sicher, dass die Anwendung auf die neuesten Informationen zurückgreift. Feedbackschleifen mit den Nutzenden der Anwendung ermöglichen, dass Datenqualität und Vorhersagen kontinuierlich verbessert werden. Die Datenlabore können hierbei unterstützen.

☐ **3.2.6 Wo sollen die Daten gespeichert werden? Gibt es langfristig ausreichend Speicherplatz?** **Umsetzende**

Der Speicherort ist abhängig vom Schutzbedarf der verwendeten Daten. Bei Bedarf müssen Daten aufgrund ihrer Sensibilität in den Netzen des Bundes gespeichert werden. Je nach Sensibilität der Daten ist neben einer lokalen Speicherung auch ein Cloud-Speicher, eine Datenbank oder ein Object Storage denkbar. In der Planung der Kapazitäten sollte das zu erwartende Datenwachstum geschätzt und eine entsprechend skalierbare Lösung gewählt werden.

☐ **3.2.7 Welche regelmäßigen Backup-Mechanismen sind vorgesehen, um einem Daten-/Wissensverlust vorzubeugen? Sind diese verteilt?** **Umsetzende**

Zu Backup-Mechanismen zählen regelmäßige, automatisierte Datenbackups mit Versionskontrolle, die an einem Offsite-Standort oder in einer Cloud vorgenommen und regelmäßig auf ihre ordnungsmäßige Funktion getestet werden, sowie ein Notfallwiederherstellungsplan. Bei Ausfall oder Zerstörung einzelner Systemkomponenten sollten verbleibende Komponenten unabhängig weiterlaufen.

☐ **3.2.8 Ist das KI-System für Menschen mit Einschränkungen angepasst und nutzbar? Ist das Front-End barrierefrei zugänglich?** **Umsetzende**

KI-Technologien bieten ein großes Potenzial für Menschen mit Einschränkungen, beispielsweise durch die Bereitstellung von Spracherkennungen, Echtzeitübersetzungen oder Computer Vision. Daher wird eine kritische Auseinandersetzung mit den Funktionalitäten hinsichtlich ihrer Nutzbarkeit für Menschen mit Einschränkungen und die aktive Einbeziehung dieser Nutzendengruppe empfohlen.

Um die Zugänglichkeit des Front-Ends zu verbessern, können beispielsweise Untertitel, Transkripte, eine Navigierbarkeit über die Tastatur, Textalternativen für Bilder und ähnliche Inhalte, Farbkontraste sowie angemessene Schriftgrößen helfen. Weitere Informationen dazu finden Sie hier und hier.

☐ **3.2.9 Wurde bereits ein Modell ausgewählt? Sind die Gründe für die Auswahl des Modells dokumentiert?** **Umsetzende**

Viele Modelle sind auf unterschiedliche Aufgaben spezialisiert. Zudem variieren Leistung, Anforderungen, Skalierbarkeit, Kosten und etwaige Datenschutzbedarfe. Aus Transparenzgründen sollte die Entscheidung für ein Modells daher gut dokumentiert sein.

☐ **3.2.10 Inwiefern sind die methodischen Entscheidungen der KI-Entwicklung dokumentiert und nachvollziehbar für interne und externe Prüfende oder Auditoren?** **Umsetzende**

Da methodische Entscheidungen einen Einfluss auf Sicherheit, Kosten, Zugänglichkeit, Abhängigkeiten, Datenverarbeitung, Skalierbarkeit und Ergebnisse haben, sollten sie in geeigneter Weise (beispielsweise in einer Entscheidungsmatrix, einem Entscheidungsprotokoll oder einem Wiki) nachvollziehbar dokumentiert werden.

☐ **3.2.11 Welche Maßnahmen sind vorgesehen, um die Möglichkeit eines Ausfallens des KI-Systems zu minimieren? Gibt es Parallelstrukturen, die eine Lauffähigkeit bei einem Systemausfall garantieren?**

Entscheidende

Umsetzende

Verschiedene Maßnahmen zur Minimierung eines Ausfallrisikos können ergriffen werden: die Implementierung von Testprozessen und Mechanismen zur Sicherstellung der frühzeitigen Fehlererkennung, die Einrichtung eines Überwachungssystems zur Identifikation von Problemen und Abweichungen sowie die Implementierung von Sicherheitsmaßnahmen zum Schutz vor Manipulation, Eindringen oder Missbrauch. Zudem sollten Parallelstrukturen auf Hard- und Softwareebene eingerichtet werden.

Mögliche Krisenszenarien sollten weiterhin in einem Notfallplan vordefiniert werden. Dieser sollte auch klarstellen, ob bei Bedarf eine Teilstilllegung möglich ist oder zwingend das ganze System abzuschalten ist. Weitere Informationen dazu finden Sie hier.

Verantwortungsvolle Nutzung und Organisation

☐ **3.3.1 Auf welche möglichen kognitiven Verzerrungen (Bias) haben Sie den Datensatz und die Ergebnisse überprüft? Welche Gegenmaßnahmen haben Sie getroffen?**

Umsetzende

Zu möglichen kognitiven Verzerrungen zählen Bias in den Trainingsdaten (z. B. eine unzureichende Repräsentanz aller Bevölkerungsgruppen), algorithmische Bias (Bevorzugung/ Benachteiligung von Gruppen aufgrund der Konstruktion des Algorithmus) oder kontextuelle Bias (falsche Schlussfolgerungen aufgrund von Verständnisproblemen der KI).

Entsprechende Gegenmaßnahmen können die Sicherstellung von Datenvielfalt, Transparenz über die verwendeten Algorithmen und Modelle sowie die Implementierung von Ethikrichtlinien für die Nutzung von KI umfassen.

4 Validierung und Evaluierung

Technische Umsetzung

☐ **4.1.1 Ist ein gut begründetes und dokumentiertes Testkonzept vorhanden, das den Umfang, die Ziele, die Methoden und die Kriterien für die Bewertung der KI-Systeme klar darlegt?**

Entscheidende

Umsetzende

Die Erstellung eines Testkonzepts ist unabdingbar, um sicherzustellen, dass die Anwendung zuverlässig, genau und robust funktioniert. Zudem ist eine Benchmark als Referenzmaßstab festzulegen, um bewerten zu können, ab wann das System als erfolgreich eingestuft wird. Die Entscheidung über Länge und Ausmaß der Pilotierung sollte transparent und vor dem Hintergrund einer Kosten-Nutzen-Abwägung getroffen werden. Weitere Informationen dazu finden Sie hier und hier.

☐ **4.1.2 Gibt es einen Plan, wie bei dem Nichterreichen der Testziele vorgegangen werden soll?**

Umsetzende

In einem ersten Schritt empfiehlt sich die Analyse der zur Schulung des Modells verwendeten Daten und eine Prüfung der Notwendigkeit, bessere und neue Daten zu sammeln. Die Überprüfung der Architektur und der Parameter des Modells können möglicherweise nahelegen, diese anzupassen und zu optimieren oder eine andere Modellstruktur zu verwenden. Weiterhin sind die Merkmale des Modells, mit deren Hilfe es Vorhersagen trifft, sowie die Trainingsprozesse zu evaluieren.

☐ **4.1.3 Haben Sie die Ergebnisse des KI-Systems auf ihre Robustheit hin überprüft?**

Umsetzende

Zur Überprüfung der Robustheit empfehlen sich Stresstests mit hoher Anfragelast oder verzerrten Eingaben und Versuche, als Nutzende/r ein Bias zu produzieren. Es sollte getestet werden, ob die gleichen Anfragen zu replizierbaren Ergebnissen führen. Das Fraunhofer-Institut für Kognitive Systeme hat hierfür das Robuscope entwickelt, das KI-Modelle online auf ihre Robustheit testet.

5 Implementierung und Integration

Verantwortungsvolle Nutzung und Organisation

☐ **5.1.1 Gibt es ein Konzept zur Überprüfung der Leistungsfähigkeit des KI-Systems im laufenden Betrieb (inklusive Benchmarking)?** Umsetzende

Die Leistung (Zuverlässigkeit, Ressourcennutzung, Geschwindigkeit) und Nutzendenzufriedenheit der KI-Lösung sollten in Echtzeit erhoben werden. Bei Fehlermeldungen sind die Administratorinnen und Administratoren sofort zu benachrichtigen.

☐ **5.1.2 Haben Sie ein Konzept zur Veränderungsbegleitung während der Umstellung auf das neue KI-System? Wie werden sich die Prozesse und Arbeitsweisen der Mitarbeitenden verändern und wie werden diese Veränderungen gesteuert, kommuniziert und ausgeglichen?** Entscheidende Umsetzende

Die Einführung neuer KI-Systeme sollte durch Veränderungsbegleitung zur Erhöhung der Akzeptanz ergänzt werden. Diese Begleitung kann die Veränderung benötigter Kompetenzen, die Beteiligung relevanter Stakeholder, die Anpassung von Strukturen und Prozessen sowie geeignete Wege zur Einbindung von Nutzenden umfassen. Konkret können darunter Trainings für Nutzende fallen, in denen die Funktionsweise erklärt und das KI-System unter Anleitung ausprobiert werden kann.

Darüber hinaus kann es Aufgabe der Veränderungsbegleitung sein, geeignete Formate und Methoden für die Einbindung von Nutzenden vor, während und nach der Entwicklung durchzuführen. Die frühzeitige Einbindung relevanter Gremien (z. B. Personalrat) ist dringend geboten.

☐ **5.1.3 Ist ein nachhaltiges Konzept zur Befähigung Ihrer Nutzenden im Umgang mit dem neuen KI-System vorhanden?** Entscheidende Umsetzende

Aufgrund möglicher Weiterentwicklungen des Systems sollten neben der initialen Befähigung von Mitarbeitenden für die Nutzung eines KI-Systems regelmäßige Weiterbildungsmöglichkeiten zur Verfügung stehen. Der Aufbau von Best Practices zur Befähigung ist ebenfalls empfehlenswert.

☐ **5.1.4 Gibt es einen Plan für die Überführung der Entwicklung in einen Regelbetrieb (u. a. Kompetenzaufbau, Protokolle, Wartung)?** Umsetzende

Beim Übergang von Entwicklung in einen Regelbetrieb gibt es vielschichtige Aspekte zu beachten, dazu gehören unter anderem:

- Befähigung der Nutzenden: Der Implementierung des KI-Systems sollte eine Befähigung der Nutzenden vorausgehen. Das Einbeziehen der Nutzenden in die Definition von Anforderungen, die Entwicklung sowie in die Pilotphase hilft, das System an ihren Bedürfnissen auszurichten und somit die Implementierung zu erleichtern.
- Frühzeitiges Festlegen von Verantwortlichkeiten: Häufig wird die Entwicklung von KI-Systemen extern ausgelagert. Eine frühzeitige Definition der Verantwortlichkeiten bei der Übergabe des KI-Systems an die Behörde in den laufenden Betrieb hilft, die Nachvollziehbarkeit zu erhöhen und technisch beziehungsweise infrastrukturelle Hürden aus dem Weg zu räumen.

- Ressourcen für Support einplanen: Insbesondere zu Beginn der Bereitstellung ist mit einem erhöhten Supportaufkommen zu rechnen. Entsprechende Ressourcen sind daher einzuplanen.
- Monitoring des Echtbetriebs einplanen: Mit dem Ziel, Fehler zu identifizieren und Modelldrifts zu vermeiden, sollten die Leistung und die Compliance mit ethischen und rechtlichen Vorgaben im Echtbetrieb überwacht werden.
- Sicherheitstests in Produktivumgebung durchführen: Bei der Inbetriebnahme im Produktivsystem kann es zu Abweichungen der Schnittstellen kommen. Deshalb sollten Sicherheitstests in der Produktivumgebung durchgeführt werden, um Risiken aufdecken zu können.
- Nutzenden-Feedback ermöglichen: Im Echtbetrieb sollte den Nutzenden die Möglichkeit eingeräumt werden, niedrigschwellig Feedback zu geben, da dieses zur Weiterentwicklung des Systems benötigt wird.

6 Wartung und Weiterentwicklung

Verantwortungsvolle Nutzung und Organisation

☐ **6.1.1 Sind die Zuständigkeiten für die nutzendenzentrierten Weiterentwicklung des KI-Systems geklärt?**

Entscheidende

Umsetzende

Neben verbindlichen Prozessen und Verfahren zum Feedback sollten Produktverantwortliche, die für die Weiterentwicklung von KI-Systemen zuständig sind, festgelegt werden. Im Zuge der Weiterentwicklung umgesetzte sowie gegebenenfalls von Nutzenden vorgeschlagene, jedoch abgelehnte Anforderungen an das KI-System sind ebenfalls zu kommunizieren und bei Bedarf zu erklären.

☐ **6.1.2 Gibt es eine Möglichkeit, das durch die Nutzung des KI-Systems erworbene Wissen innerhalb der Bundesverwaltung zu teilen und damit anderen Ressorts die Nachnutzung des KI-Systems zu ermöglichen?**

Umsetzende

Zum Austausch von Erfahrungswerten und Wissen über KI-Systeme innerhalb der Bundesverwaltung eignen sich etablierte Plattformen wie der „Marktplatz der KI-Möglichkeiten“.