VORLAGE

KI-Richtlinien

Dokumentenkontrolle	
Verantwortlicher für die Richtlinie:	
Vorbereitet von:	
Genehmigt durch:	
Datum der Genehmigung:	
Datum der nächsten Überprüfung:	

Versionsgeschichte		
Versionsnummer	**Zusammenfassung der Änderungen**	**Datum**

1. **Einführung**

Ziel dieser KI-Richtlinie ist es, den Ansatz von Name des Unternehmens] in Bezug auf künstliche Intelligenz, einschließlich der Konzeption, der Entwicklung, des Einsatzes, der Nutzung und der Beschaffung von KI-Systemen, darzustellen.

[Name des Unternehmens] erkennt an, dass KI-Systeme einen enormen Wert schaffen können, gleichzeitig aber auch Datenschutz-, ethische, rechtliche und Informationssicherheitserwägungen mit sich bringen, die bei allen Projekten und Initiativen, die KI betreffen, berücksichtigt werden müssen.

Diese Richtlinie schafft daher einen Rahmen, der sicherstellen soll, dass KI-Technologien in einer Weise eingesetzt werden, die ethisch vertretbar und transparent ist und den einschlägigen Vorschriften entspricht.

2. **Umfang**

Diese Richtlinie zur KI gilt für alle Projekte und Initiativen, die KI-Systeme betreffen, einschließlich, aber nicht beschränkt auf:

- Design, Entwicklung, Testungen sowie Einsatz und Überwachung von KI-Algorithmen und -Systemen, entweder intern oder durch Drittanbieter.
- Die Nutzung von KI-Systemen durch Mitarbeiter, Auftragnehmer oder andere Stellen, die im Namen von Name des Unternehmens] handeln.
- Die Beschaffung von KI-Systemen von Drittanbietern.

3. **Definitionen**

Künstliche Intelligenz (KI): Bezieht sich auf Computersysteme, die die menschliche Intelligenz simulieren und Aufgaben wie das Lernen aus Daten, das Erkennen von Mustern, das Treffen von Entscheidungen und das Lösen von Problemen übernehmen. Sie umfasst maschinelles Lernen, natürliche Sprachverarbeitung und Robotik.

KI-Algorithmus: Ein schrittweiser Satz von Anweisungen oder ein Rechenverfahren, das zur Lösung eines bestimmten Problems oder zur Durchführung einer bestimmten Aufgabe im Bereich der künstlichen Intelligenz entwickelt wurde. Diese Algorithmen werden verwendet, um Daten zu verarbeiten, Vorhersagen zu treffen oder Maßnahmen auf der Grundlage von Eingabedaten und Regeln zu ergreifen, die von Programmierern definiert oder durch maschinelles Lernen aus Daten gelernt wurden.

KI-System: Ein System, das sowohl die KI-Algorithmen als auch die für die Umsetzung der Fähigkeiten der künstlichen Intelligenz erforderliche Hardware- oder Software-Infrastruktur umfasst. Es umfasst die Kombination von KI-Algorithmen, Datenspeicherung, Verarbeitungseinheiten und häufig auch Schnittstellen für die Ein- und Ausgabe.

KI-Projekt: Jede Geschäftsinitiative, die den Einsatz von KI-Systemen beinhaltet, einschließlich der Entwicklung, Nutzung oder Beschaffung von Lösungen, die KI-Systeme verwenden.

Generative Künstliche Intelligenz (GenKI): Ein Teilbereich der künstlichen Intelligenz, der sich auf die eigenständige Erstellung von Inhalten wie Bildern, Text oder Musik konzentriert. Sie setzt Algorithmen ein, um in kürzester Zeit

menschenähnliche kreative Ergebnisse zu erzeugen.

KI-Lebenszyklus: Bezeichnet die Phasen der Entwicklung und Nutzung eines KI-Systems, von der Idee bis zum Einsatz und zur kontinuierlichen Verbesserung. Diese Phasen sind wie folgt definiert:

- **Design & Steuerung:** In dieser Phase werden die Ziele und der Umfang des KI-Systems definiert sowie die Gesamtarchitektur und die Anforderungen an das KI-System umrissen.
- **Datenerfassung:** In dieser Phase werden Daten gesammelt und aufbereitet, um die Lern- und Entscheidungsprozesse des KI-Systems zu unterstützen.
- **Modellentwicklung:** Das KI-System wird entwickelt, gebaut, trainiert und getestet, um sicherzustellen, dass diese neue Ergebnisse erzeugen kann.
- **Bereitstellung & Überwachung:** Das trainierte KI-System wird in der realen Umgebung eingesetzt, und seine Leistung und sein Verhalten werden kontinuierlich überwacht, um die volle Funktionalität sicherzustellen und um eventuelle Verbesserungsmöglichkeiten zu ermitteln.

KI-Grundsätze: Beziehen sich auf die grundlegenden Richtlinien und Werte, die einen Rahmen für die Entwicklung und Nutzung von KI-Systemen bilden. Diese Grundsätze zielen darauf ab, ethische und vertrauenswürdige KI-Praktiken zu etablieren und gleichzeitig einen menschenzentrierten Ansatz für die KI-Technologie zu fördern und sind in Abschnitt 4 dieser Richtlinie niedergelegt.

Datenschutzgesetze: Geltende Datenschutzgesetze, einschließlich, aber nicht beschränkt auf die EU-DSGVO und die britische Version der Verordnung (UK GDPR).

Systemverantwortliche(r): Die Person(en), die in erster Linie für die Entwicklung und/oder den Einsatz eines bestimmten KI-Systems verantwortlich und rechenschaftspflichtig ist/sind.

Verantwortliches KI-Überwachungsteam: Das Team, das für die Überwachung der Entwicklung, des Einsatzes und der Nutzung von KI-Systemen innerhalb der Organisation verantwortlich ist.

4. **Rollen und Zuständigkeiten**

4.1 Der Datenschutzbeauftragte (DSB) berät alle relevanten Interessengruppen und insbesondere den Systemverantwortlichen bezüglich der Übereinstimmung zwischen den eingesetzten KI-Systemen mit den geltenden Datenschutzvorschriften sowie den Vorgaben der Datenverwaltung. Letztere sind in Punkt 5.4. dieser KI-Richtlinie festgehalten. Zudem ermittelt der Datenschutzbeauftragte diesbezügliche Risiken und dazugehörige Abhilfemaßnahmen.

4.2. Alle anderen Mitglieder des KI-Überwachungsteams sind dafür verantwortlich, die Systemverantwortlichen und andere relevanten Beteiligten hinsichtlich der Konformität von KI-Systemen mit den in Teil 5 dargelegten KI-Grundsätzen zu beraten, und zwar in Übereinstimmung mit dem Fachgebiet des jeweiligen Teammitglieds. Zudem ermitteln auch sie, wiederum innerhalb ihres jeweiligen Fachgebietes, Risiken und Maßnahmen zur Risikominderung.

4.3. Die Aufgaben der Systemverantwortlichen sind wie folgt:

- Gewährleistung der Umsetzung geeigneter technischer und organisatorischer Maßnahmen, die darauf abzielen, die KI-Systeme mit Unterstützung des zuständigen KI-Überwachungsteams an die in Ziffer 5 dieser Richtlinie genannten KI-Grundsätze anzupassen und
- Sicherstellung einer effektiven Zusammenarbeit mit dem zuständigen KI-Überwachungsteam, wenn dies in Übereinstimmung mit Teil 6 dieser Richtlinie erforderlich ist.

4.4. Alle an der Entwicklung bzw. dem Einsatz von KI-Systemen beteiligten Akteure sind dafür verantwortlich, die Systemverantwortlichen bei der Umsetzung geeigneter technischer und organisatorischer Maßnahmen zu unterstützen. Diese Maßnahmen haben das Ziel die KI-Systeme mit den in Teil 5 dieser Richtlinie dargelegten KI-Grundsätzen in Einklang zu bringen.

5. **KI-Prinzipien**

5.1. Rechenschaftspflicht

5.1.1. Die Einhaltung dieser Grundsätze und Anforderungen ist während des gesamten KI-Lebenszyklus zu berücksichtigen und zu überprüfen. Vor dem Beginn jeder Phase des KI-Lebenszyklus muss der Systemverantwortliche mit dem zuständigen KI-Überwachungsteam zusammenarbeiten, um neue Risiken zu ermitteln und ggf. zu mindern, bevor er fortfährt.

5.1.2. Aufzeichnungen und Unterlagen, die die Einhaltung dieser Grundsätze und Anforderungen belegen sind aufzubewahren. Ist eine sogenannte Datenschutz-Folgenabschätzungen (DSFA) erforderlich, ist diese ebenso zu dokumentieren. Wurden die KI-Systeme intern entwickelt sind zudem alle dazugehörigen technischen Unterlagen aufzubewahren.

5.2. Menschliches Handeln und Kontrolle

Der Grundsatz des menschlichen Handelns und der Kontrolle bedeutet, dass KI-Systeme als ein Werkzeug entwickelt werden, das Menschen dient und die Menschenwürde sowie persönliche Autonomie respektiert.

Alle KI-Systeme sollten Mechanismen enthalten, die die menschliche Aufsicht und Intervention ermöglichen.

5.2.1**.** Art und Umfang der menschlichen Aufsichts- und Eingriffsmechanismen müssen auf die Art des KI-Systems und das damit verbundene Risiko in Bezug auf das Datenschutzrecht und die KI-Grundsätze abgestimmt sein. In diesem Sinne muss der Systemverantwortliche den Rat des zuständigen KI-Überwachungsteams hinsichtlich geeigneter menschlicher Aufsichts- und Eingriffsmechanismen im Einklang mit den in Abschnitt 6.2 dieser Richtlinie beschriebenen Verfahren einholen.

5.2.2. Der Systemverantwortliche und das zuständige KI-Überwachungsteam müssen zumindest prüfen, ob das KI-System dazu verwendet werden könnte, automatisierte Entscheidungen zu treffen, die erhebliche (z. B. finanzielle oder rechtliche) Auswirkungen auf Einzelpersonen haben können. Dazu gehört auch die Ermittlung von notwenigen, um eine angemessene Aufsicht und Abhilfe zu gewährleisten.

5.3. Technische Robustheit und Sicherheit

5.3.1. Der Grundsatz der technischen Robustheit und Sicherheit bedeutet, dass KI-Systeme so entwickelt und eingesetzt werden, dass unbeabsichtigte oder unerwartete Schäden minimiert werden. Zu diesem Zweck sollten die Systemeigentümer mit Unterstützung und Beratung durch das zuständige KI-Überwachungsteam sicherstellen, dass verhältnismäßige, risikobasierte Maßnahmen zur Optimierung der Modellrobustheit, Zuverlässigkeit, Genauigkeit und Informationssicherheit getroffen werden.

Diese Maßnahmen werden:

- In der Design- und Governance-Phase des KI-Lebenszyklus geprüft
- in der Phase der Modellentwicklung, Schulung und Testung implementiert und getestet und
- in der Phase "Einsatz und Überwachung" überwacht.

5.3.2. Alle Mitarbeiter, die an der Entwicklung, dem Einsatz oder der Nutzung beteiligt sind, sind dafür verantwortlich, dass die Entwicklung, der Einsatz und die Nutzung von KI-Systemen in Übereinstimmung mit den internen Richtlinien und Maßnahmen zur Informationssicherheit von [Name des Unternehmens] erfolgt.

5.3.3. Vor dem Einsatz von KI-Systemen muss eine Überprüfung und Bewertung durchgeführt werden, um die direkten oder indirekten Auswirkungen auf die bestehenden organisatorischen und technischen Kontrollen für die Informationssicherheit zu bestimmen sowie alle vernünftigerweise vorhersehbaren Informations- und Sicherheitsrisiken zu ermitteln, und zwar in Übereinstimmung mit den in Abschnitt 6.2 dieser Richtlinie beschriebenen Verfahren.

5.3.4. Vor dem Einsatz von KI-Systemen sollten angemessene Maßnahmen zur Minderung solcher Risiken ergriffen werden, z. B. Schutzvorkehrungen gegen gegnerische Angriffe, unerwartete Eingabedaten oder Manipulationsversuche. Solche Maßnahmen sollten von Fall zu Fall bewertet und dokumentiert werden, wobei der Rat des Informationssicherheitsexperten des Unternehmens zu berücksichtigen ist.

5.4. Datenschutz und Datenverwaltung

5.4.1. Der Grundsatz Datenschutzes und der Datenverwaltung besagt, dass KI-Systeme unter Einhaltung der Datenschutzgesetze bei der Verarbeitung personenbezogener Daten entwickelt und eingesetzt werden müssen.

5.4.2. Demnach müssen alle KI-Systeme, die personenbezogene Daten verarbeiten, die gesetzlichen Vorgaben des Datenschutzes erfüllen, einschließlich, aber nicht beschränkt auf die folgenden Anforderungen:

- Für alle Verarbeitungen personenbezogener Daten durch oder im Zusammenhang mit einem KI-System muss eine Rechtsgrundlage gemäß Artikel 6 der DSGVO ermittelt und dokumentiert werden. Werden darüber hinaus sensible Daten verarbeitet ist die Erfüllung einer geeignete Bedingung gemäß Artikel 9 der DSGVO ebenso zu dokumentieren.
- Die internen Aufzeichnungen über die Verarbeitungstätigkeiten müssen entsprechend den im KI-System vorgesehenen Verarbeitungen aktualisiert werden.
- Einschlägige Richtlinien und Hinweise zum Datenschutz müssen aktualisiert werden, um klare Informationen über die Verarbeitung personenbezogener Daten im Zusammenhang mit dem KI-System zu liefern.

5.4.3. Der Systemverantwortliche holt zur Sicherstellung der Einhaltung der Bestimmungen in Teil 5.4.1 den Rat des Datenschutzbeauftragten ein. Das einzuhaltende Verfahren ist in Teil 6.2 dieser Richtlinie beschrieben.

5.5. Vielfalt, Nichtdiskriminierung und Fairness

5.5.1 Der Grundsatz der Vielfalt, Nichtdiskriminierung und Fairness bedeutet, dass KI-Algorithmen so entwickelt und

eingesetzt werden, dass sie einen gleichberechtigten Zugang fördern und Diskriminierung und unfaire Vorurteile aufgrund von Ethnie, Geschlecht, Alter oder anderen geschützten Merkmalen vermeiden.

5.5.2. Bei der Entwicklung von KI-Systemen sollte eine Bewertung durchgeführt werden, um die nach vernünftigem Ermessen vorhersehbaren Risiken ungerechter Ergebnisse für die betroffenen Interessengruppen zu ermitteln. Das entsprechende Verfahren ist in Teil 6.2 dieser Richtlinie festgelegt.

5.5.3. Im Anschluss an Teil 5.5.2. sollten in Fällen, in denen die Gefahr ungerechter Ergebnisse für die betroffenen Interessengruppen besteht, verhältnismäßige Maßnahmen und Prüfverfahren eingeführt werden, um diese Risiken zu mindern.

5.5.4. KI-Systeme sollten regelmäßig überwacht werden, um Vielfalt, Nichtdiskriminierung und Fairness zu gewährleisten. Werden in einem KI-System unfaire Ergebnisse festgestellt, muss der Systemverantwortliche sicherstellen, dass die erforderlichen Abhilfemaßnahmen ergriffen werden.

5.6. Transparenz und Verständlichkeit

5.6.1. Der Grundsatz der Transparenz und Erklärbarkeit bedeutet, dass die Ergebnisse von KI-Systemen klar und verständlich erläutert werden.

5.6.2. Bei der Entwicklung von KI-Systemen stellt der Systemverantwortliche sicher, dass eine Dokumentation des KI-Systems mit detaillierten Informationen über Datenquellen, verwendete Algorithmen und Entscheidungsprozesse geführt wird.

5.6.3. Personen, die direkt mit KI-Systemen zu tun haben, müssen über die Fähigkeiten und Grenzen solcher KI-Systeme informiert werden.

5.6.4. Personen, die unmittelbar von der Entwicklung und dem Einsatz von KI-Systemen betroffen sind, müssen über ihre Rechte gem. den gesetzlichen Bestimmungen informiert werden, bspw. im Rahmen des Datenschutzrechts.

6. **Verwaltung und Risikomanagement**

6.1. Verantwortliches KI-Überwachungsteam

6.1.1. Dem KI-Überwachungsteam gehören zumindest der Datenschutzbeauftragte sowie Experten für KI-Ethik, Informationssicherheit und Rechtskonformität an.

6.1.2. Zu den Mitgliedern des verantwortlichen KI-Überwachungsteams können sowohl interne als auch externe Personen gehören.

6.1.3. Die Mitglieder des zuständigen KI-Überwachungsteams sind in Anhang 1 dieser Richtlinie aufgeführt.

6.2. Beurteilungen und Risikomanagement

6.2.1. Die Systemverantwortlichen stellen sicher, dass alle vorgeschlagenen Verwendungszwecke neuer KI-Systeme einer Screening-Bewertung unterzogen werden, die dem zuständigen KI-Überwachungsteam vor Beginn des KI-Projekts zur Prüfung vorgelegt wird.

6.2.2. Stellt der Datenschutzbeauftragte fest, dass eine Datenschutz-Folgenabschätzung durchgeführt werden muss, wird eine solche vor jeder Datenerhebung oder sonstigen Datenverarbeitung im Zusammenhang mit dem KI-Projekt durchgeführt.

6.2.3. Das verantwortliche KI-Überwachungsteam ist für die Ermittlung und Koordinierung zusätzlicher Bewertungen und die Beratung über Korrekturmaßnahmen zur wirksamen Steuerung und Minderung bestehender Risiken zuständig.

6.2.4. Alle relevanten Prüfungen sind während der gesamten Dauer des KI-Lebenszyklus durchzuführen und müssen mindestens vor Beginn jeder Phase des KI-Lebenszyklus erneut durchgeführt werden. Wie in Abschnitt 4.3 festgelegt, werden solche Prüfungen vom Systemverantwortlichen koordiniert und beteiligt zumindest die Mitglieder des KI-Überwachungsteams.

6.3. Schulung und Sensibilisierung

6.3.1. [Name des Unternehmens] wird Schulungs- und Sensibilisierungsprogramme anbieten, um die Mitarbeiter über den ethischen und verantwortungsvollen Einsatz von KI aufzuklären.

6.4. Lieferanten von Drittanbieter

6.4.1. Bei der Beschaffung von KI-Systemen oder von Produkten bzw. Dienstleistungen, die KI-Systeme nutzen, müssen die Verantwortlichen sicherstellen, dass alle internen Richtlinien und Verfahren bezüglich Beschaffung und Sorgfaltspflicht des Verkäufers im jeweiligen Beschaffungsprozess befolgt werden.

6.4.2. Bei der Beschaffung von KI-Systemen von Drittanbietern wird Name des Unternehmens] Maßnahmen ergreifen, um

sicherzustellen, dass diese KI-Systeme den in dieser Richtlinie dargelegten KI-Grundsätze im Wesentlichen entsprechen.

6.4.3. Bei der Beschaffung über externe Dritte von KI-Systemen oder von Produkten bzw. Dienstleistungen, die KI-Systeme nutzen, müssen die Verantwortlichen vor dem Vertragsabschluss von diesen Dritten verlangen, dass sie eine Anbieterbewertung ausfüllen. Diese Bewertung dient der Sammlung von Informationen über die KI-Praktiken des Dritten.

6.4.4. Die ausgefüllten Bewertungen der Anbieter werden dem zuständigen KI-Überwachungsteam zusammen mit der Screening-Bewertung vorgelegt, bevor ein Vertrag mit Drittanbietern geschlossen wird.
Wenn Drittanbieter die Anbieterbewertungen nicht ausfüllen, nutzt das zuständige KI-Überwachungsteam andere verfügbaren Informationen, um die Risiken des KI-Systems des Drittanbieters zu identifizieren, zu analysieren und zu den KI-Grundsätzen zu beraten.

7. **Nutzung der GenKI-Tools durch Mitarbeiter**

7.1. Mitarbeiter von [Name des Unternehmens] sind berechtigt, GenKI-Tools zur Unterstützung und Optimierung ihrer Arbeit zu nutzen, vorbehaltlich der Bedingungen und Einschränkungen, die in der „Generative KI-Richtlinie zur zulässigen Nutzung“ festgelegt sind.

7.2. Wenn Mitarbeiter sich nicht sicher sind, ob ihre voraussichtliche Nutzung von GenKI-Tools, die in der entsprechenden Richtlinie der zulässigen Nutzung von generativer KI festgelegten Bedingungen erfüllt sind, sollten sie sich an das zuständige KI-Überwachungsteam wenden.

8. **Überprüfung der Richtlinie**

8.1. Diese Richtlinie ist jährlich zu überprüfen. Jede Überprüfung wird vom Verantwortlichen der Richtlinie koordiniert und bezieht die Systemverantwortlichen und das zuständige KI-Überwachungsteam ein.

8.2. Das nächste Überprüfungsdatum wird im Abschnitt über die Dokumentenkontrolle in dieser Richtlinie festgelegt.

8.3. Die Ergebnisse jeder Überprüfung und die sich daraus ergebenden Änderungen dieser Richtlinie bedürfen der Genehmigung durch den Verantwortlichen der Richtlinie und werden in der Versionsgeschichte dieser Richtlinie festgehalten.

Anhang 1: Verantwortliches KI-Überwachungsteam

Teammitglied(er) und Rolle(n)	**KI-Fachwissen Bereich**
[DataGuard], Datenschutzbeauftragter	Datenschutz und Privacy