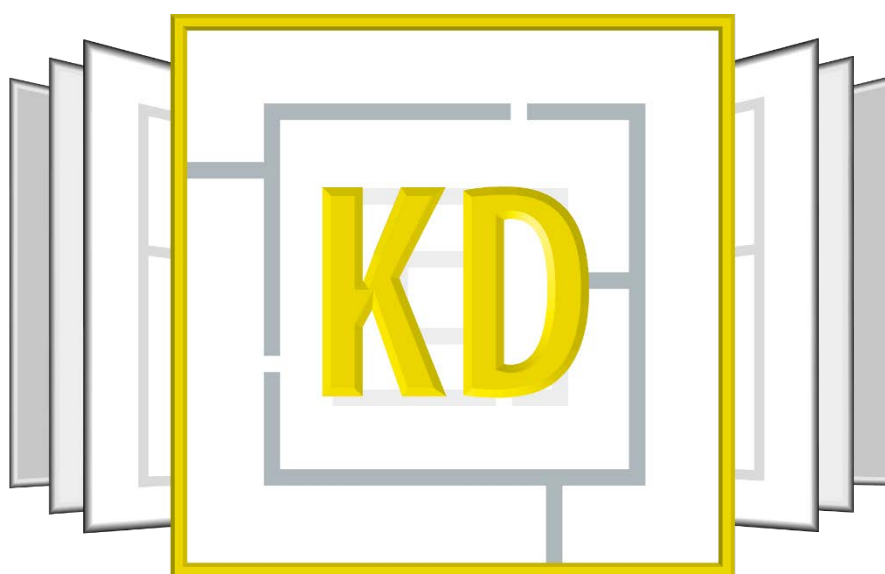




Künstliche Intelligenz

K-9000/102



Konzeptionelle Dokumentenlandschaft

ÖFFENTLICHE FASSUNG

Diese öffentliche Fassung basiert weitgehend auf der eingestuften Fassung.



Strategisch-politische
Dokumente



Allgemeine
Regelungen



Druckschriften



Technische Regelungen



Regelungsnahe
Dokumente

Detailinformationen

Zweck der Regelung:	Beitrag zur strategischen Steuerung der Entwicklung, Einführung und Nutzung Künstlicher Intelligenz im Geschäftsbereich des Bundesministeriums der Verteidigung
Geltungsbereich:	Geschäftsbereich des Bundesministeriums der Verteidigung
Datum Gültigkeitsbeginn:	09.04.2026
Herausgebende Stelle:	BMVg
Regelungsnummer, Version:	K-9000/102, Öffentliche Fassung
Aktenzeichen:	62-00-03
Gebilligt durch:	Abteilungsleiter Innovation und Cyber

Mögliche Kennzeichnungen

Ä	Änderungen zur vorherigen Veröffentlichung	B	Berichtspflichten
!	Besonders wichtige Wörter, Zeilen oder Abschnitte	E	Abweichende Vorgaben für den Einsatz
Y	Befehle im Sinne des § 2 Nr. 2 WStG	S	Sicherheitsbestimmungen

Inhaltsverzeichnis

1	Einordnung	5
1.1	Begriffsbestimmungen	5
1.2	Zielsetzung und strategische Grundlagen	6
1.3	Voraussetzungen und Einbettung	7
2	Rahmenbedingungen	8
2.1	Politischer Rahmen	8
2.2	Rechtlicher Rahmen	9
2.3	Ethischer Rahmen	12
2.4	Internationaler Rahmen	14
3	Handlungsfelder	16
3.1	Handlungsfeld Ausbildung und Qualifizierung	16
3.1.1	Allgemeine Ausbildung und Qualifizierung	16
3.1.2	Erweiterte Ausbildung und Qualifizierung	17
3.1.3	Ausbildung und Qualifizierung von Führungskräften	18
3.1.4	Aufgabenbezogene Ausbildung und Qualifizierung	19
3.1.5	Künstliche Intelligenz in der Ausbildung und Qualifizierung	19
3.2	Handlungsfeld Fähigkeiten	20
3.2.1	Vorgaben für die Fähigkeitsentwicklung der Bereiche	20
3.2.2	Steuerung, Planung Teilportfolio Cyber/IT und Beschaffung	22
3.2.3	KI-Infrastruktur	22
3.3	Handlungsfeld Verantwortungsvolle Nutzung von Künstlicher Intelligenz	23
3.3.1	Risikomanagement für Künstliche Intelligenz	24
3.3.2	Transparenz, Rückverfolgbarkeit, Dokumentation	25
3.3.3	Trainieren, Verifizieren, Validieren und Testen	28
3.3.4	Konformitätsprüfung, Genehmigung zur Nutzung und Inbetriebnahme	30
3.3.5	Menschliche Kontrolle	31
3.3.6	Zuverlässigkeit, Robustheit, Sicherheit und Datenschutz	32
4	KI-Governance und Umsetzungsschritte	34
4.1	KI-Governance im Geschäftsbereich des BMVg	34
5	Anlagen	37
5.1	Bereichsausnahme von der EU KI-Verordnung	38
5.1.1	Stellungnahme BMVg vom 30. Januar 2025 zur Reichweite der Bereichsausnahme	38
5.2	Erläuterung zur Einführung eines Risikomanagements Künstlicher Intelligenz im Geschäftsbereich des BMVg	45
5.2.1	Risikoidentifizierung, -bewertung und -steuerung	45
5.2.2	Risiko-Definitionen	47
5.3	Abkürzungsverzeichnis	51

1 Einordnung

101. Dieses Konzeptionelle Dokument (KD) bildet den Rahmen für die rechtskonforme, verantwortungsvolle, sichere und zuverlässige Nutzung (einschließlich Erprobung) von Künstlicher Intelligenz (KI) im Geschäftsbereich des Bundesministeriums der Verteidigung (GB BMVg) und zeigt in Handlungsfeldern die notwendigen Handlungsbedarfe auf.

102. Da KI zunehmend in (IT-)Systemen Einzug hält, ist es wichtig, bei der Anwendung dieses KD und den hieraus abzuleitenden Handlungsbedarfen und Regelwerken einen pragmatischen und dennoch rechtssicheren Ansatz zu verfolgen. Dieses KD richtet sich an alle Planungskategorien im Planungsprozess. KI ist – ebenso wie andere informationstechnische Fähigkeiten – ein Instrument zur sachgerechten Erfüllung der fachlichen Aufgaben sowie zur Weiterentwicklung. Sie trägt auf diesem Wege zur Modernisierung und Zukunftsfähigkeit des GB bei. Ihre Nutzung ist nur dann nachhaltig wirksam, wenn sie in die bestehenden Geschäftsprozesse, Verfahrensabläufe und Verantwortungsstrukturen eingebettet wird. Der Einsatz von KI steht daher in einem unmittelbaren funktionalen Zusammenhang mit der jeweiligen Fachaufgabe und den organisatorischen Rahmenbedingungen.

1.1 Begriffsbestimmungen

103. Eine allgemein akzeptierte Definition für KI existiert gegenwärtig nicht; gleiches gilt bereits für den Begriff Intelligenz. Dies ist auch eine Folge der nach wie vor schnellen Entwicklung im Themengebiet KI. Die KI-Strategie der Bundesregierung¹ sieht von einer Definition ab. Weder die EU² KI-Verordnung (EU KI-VO)³ noch die NATO⁴ KI-Strategie führen eine Definition ein.

104. Die ISO/IEC 2382:2015 „Information technology – Vocabulary“ definiert KI wie folgt: „Künstliche Intelligenz ist ein Teilgebiet der Informatik, das sich mit der Entwicklung von Datenverarbeitungssystemen befasst, die Funktionen ausführen, die normalerweise mit menschlicher Intelligenz in Verbindung gebracht werden, wie z. B. logisches Denken, Lernen und Selbstverbesserung“.⁵

105. Gesetzlich definiert und damit bindend und relevant für den GB BMVg ist der Begriff „**KI-System**“ in der EU KI-VO⁶: „Ein „KI-System“ ist ein maschinengestütztes System, das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist und das nach seiner Betriebsaufnahme anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele

¹ Strategie Künstliche Intelligenz der Bundesregierung November 2018 & Fortschreibung Dezember 2020.

² Europäische Union.

³ Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz).

⁴ North Atlantic Treaty Organization.

⁵ ISO/IEC 2382-1:2015: „Artificial Intelligence (AI) is a branch of computer science devoted to developing data processing systems that perform functions normally associated with human intelligence, such as reasoning, learning, and self-improvement.“.

⁶ EU KI-VO, Artikel 3 Absatz 1.

ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die die physische oder virtuelle Umgebung beeinflussen können.“ Auf Ebene der EU erfolgen derzeit Arbeiten zur Auslegung der Begriffe in der Definition.

106. Im allgemeinen Sprachgebrauch (und mitunter auch in weniger formalen technischen Kontexten) werden zur Beschreibung der Funktionen und Fähigkeiten von KI-Systemen häufig Begriffe verwendet, die ursprünglich dem menschlichen Erleben entstammen (z. B. „verstehen“, „lernen“, „sehen“, „denken“). Dies dient der Vereinfachung komplexer technischer Prozesse und erleichtert die intuitive Erfassung der Systemleistung. Aus technischer Sicht handelt es sich bei diesen Vorgängen ausschließlich um die Verarbeitung von Daten mittels komplexer Algorithmen und statistischer Modelle, die keine Form von Bewusstsein, Absicht oder emotionalem Erleben im menschlichen Sinne beinhalten. Die Verwendung solcher Begriffe ist zweckmäßig, sollte aber stets unter dem Vorbehalt der technischen Realität betrachtet werden, um Missverständnisse über die tatsächlichen Fähigkeiten und Grenzen der KI zu vermeiden.

1.2 Zielsetzung und strategische Grundlagen

107. KI-Systeme können die Reaktionsfähigkeit in Bezug auf die Beherrschbarkeit sowie die Nutzbarmachung großer Datenmengen verbessern und so einen operationellen Mehrwert sowohl für die Streitkräfte, in Form von Informations-, Führungs- und Wirkungsüberlegenheit, als auch innerhalb der Wehrverwaltung und Regierungsarbeit erzeugen. Die Nutzung von KI-Systemen kann zu neuen Fähigkeiten führen, vorhandene Fähigkeiten ergänzen, aber auch bisherige Fähigkeiten ersetzen. Mit dem Einsatz von KI werden Effizienzsteigerungen, Effektivitätsgewinne und neue Fähigkeiten verbunden sein, die sich auf einem zukünftigen digitalisierten Gefechtsfeld in allen Dimensionen auswirken und die Auftragserfüllung der Bundeswehr (Bw) beeinflussen werden. Darüber hinaus kann durch die Ad-hoc Rekombination von KI und Waffensystemen (WaSys) der Ansatz eines „Mosaic Warfare“⁷ verwirklicht und im Sinne von Multi-Domain Operations (MDO) die Entscheidungs- und Wirkungsüberlegenheit eigener Kräfte durch die gezielte Überforderung der gegnerischen Führung ermöglicht werden. Die KI wird damit absehbar entscheidende Auswirkungen auf die Bereitstellung einsatzbereiter und zukunftsfähiger Kräfte und deren Siegfähigkeit haben, zumal KI-Systeme bereits heute das Gefechtsfeld prägen.

108. Der GB BMVg verfolgt mit dem Einsatz von KI auch eine deutliche Optimierung des Stabs- und Verwaltungshandeln (StVwH), der Geschäftsprozesse und Fachverfahren sowie eine wirkungsvolle Unterstützung der Angehörigen in ihrer täglichen Arbeit. Die integrierte KI-Unterstützung muss

⁷ Militärisches Konzept, bei dem viele kleine, spezialisierte und vernetzte Einheiten flexibel zusammenwirken, um eine gemeinsame Wirkung zu erzielen. Es ersetzt traditionelle, zentrale Plattformen durch modulare, verteilte Systeme, die durch Autonomie, Redundanz und schnelle Anpassbarkeit gekennzeichnet sind.

beim StVwH mitgedacht und nutzbar gemacht werden.⁸ KI kann eine wertvolle Unterstützung für Menschen mit Einschränkungen sein. KI kann damit die Teilhabe – im Rahmen der Inklusion weiter verbessern.

109. Die KI ist ein wesentlicher Treiber der Digitalen Transformation des GB BMVg. Vision, Mission und Ziele der Digitalen Transformation gelten daher auch für die Nutzung von KI in der Bw. Hieraus leiten sich analog zur Digitalen Transformation unmittelbar drei Ziele in Bezug auf KI ab:

- das Erhöhen der Durchsetzungsfähigkeit der Streitkräfte auf dem digitalisierten Gefechtsfeld,
- neuartiger, effizienterer Umgang mit Daten, Informationen sowie deren Verarbeitung in Prozessen und im StVwH und
- die Steigerung der Fähigkeit der KI im gesamten GB BMVg nutzbar zu machen.

110.

1.3 Voraussetzungen und Einbettung

111. Der Erfolg der Nutzung von KI im GB BMVg steht auch im Zusammenhang mit der Quantität, Qualität und Integrität der Daten, mit denen die KI-Systeme lernen bzw. umgehen. Somit sind qualitätsgesicherte Daten und deren Verfügbarkeit auch über den Einsatz von KI hinaus eine Ressource von strategischer Bedeutung und sind gemäß Vorgaben der zentralen Data Governance im operationellen Datenmanagement, der Datendomänencluster und Datendomänen des GB BMVg vertrauenswürdig und verlässlich zu nutzen (siehe Abschnitte 3.3.6.4 und 3.3.6.5). Dies schließt explizit die Trainingsdaten für KI-Systeme und die Vertrauenswürdigkeit von KI-Modellen ein.

112. Die KI ist im Gesamtkontext der Digitalisierung des GB BMVg zu betrachten. Der strategisch-politische Rahmen wird v. a. durch die Nationale Sicherheitsstrategie⁹, die KI-Strategie der Bundesregierung, die Verteidigungspolitischen Richtlinien¹⁰ und weiteren Regelwerken des GB BMVg gesetzt.

113. Querschnittlich nutzbare KI-Systeme kommen in diversen Anwendungsfeldern zum Einsatz und werden bei Bedarf rasch weiterentwickelt sowie umfassend in der gesamten Organisation ausgerollt. Auf diese Weise können Aktualisierungen zeitnah umgesetzt und notwendige Neuerungen schnell in allen relevanten Bereichen verfügbar gemacht werden. Software Defined Defence (SDD) spielt eine zentrale Rolle, um die benötigten Werkzeuge für den Einsatz und die Etablierung von KI-Systemen vom Rechenzentrum bis hin zu den Plattformen bereitzustellen.

114. Bei der Anwendung der hier beschriebenen Rahmenbedingungen und Vorgaben ist handlungsleitend, eine verantwortungsvolle Nutzung von KI zu ermöglichen und nicht wesentlich einzu-

⁸ Die Einführung von IT-Unterstützung mit KI-Komponenten im Sinne des § 9 E-Government-Gesetz (EGovG) setzt in besonderem Maße ein abgestimmtes Zusammenwirken von Fachaufgaben, Organisation und technischer Umsetzung voraus.

⁹ SD „Nationale Sicherheitsstrategie“.

¹⁰ SD „Verteidigungspolitische Richtlinien 2023“.

schränken oder zu verzögern. Verfolgt wird ein abwägender Risikomanagementansatz (siehe Nr. 342). Eine Weiterentwicklung der Bw durch den Einsatz von Digitalisierung/KI ist im Rahmen der jeweiligen Fachaufgabe stets mit zu denken, entsprechende Forderungen sind sachgerecht in das Bedarfs- und Anforderungsmanagement einzubringen.

2 Rahmenbedingungen

2.1 Politischer Rahmen

201. Der Diskurs zum militärischen Anwendungsbereich und der konkreten Nutzung von KI durch den GB BMVg ist eingebettet in den gesamtgesellschaftlichen Diskurs zum Themenkomplex KI und den mit der Anwendung verbundenen Entwicklungsperspektiven und regulatorischen Anforderungen.

202. Im gesamtgesellschaftlichen Diskurs wird der Einsatz von KI in der Bw häufig fälschlicherweise mit letalen autonomen Waffensystemen (LAWS)¹¹ gleichgesetzt. Um Missverständnisse zu vermeiden und aufklärend zu wirken, gilt es für den GB BMVg, den gesellschaftlichen Diskurs mitzugestalten und so die Ziele des Einsatzes von KI in der Bw zu erklären sowie den Mehrwert und die Notwendigkeit des Einsatzes von KI für die Umsetzung der Auftragserfüllung der Bw herauszustellen. Deutschland (DEU) beteiligt sich an internationalen Initiativen zur Regulierung bzw. Ächtung des Einsatzes von LAWS wie dem Waffenübereinkommen der Vereinten Nationen (VN), die die Bw umsetzen muss (siehe Abschnitt 3.3.5).

203. Das Einhalten unserer Wertestandards und des Völkerrechts, insbesondere des Humanitären Völkerrechts, bilden die normative Grundlage in der Entwicklung, Einführung und Nutzung von KI im GB BMVg. Festzustellen ist, dass gegnerische Akteure die Erprobung und den Einsatz von KI außerhalb unserer Moralvorstellungen und Rechtsauffassungen intensiv verfolgen. Eigene Kräfte können auf gegnerische Kräfte treffen, die KI-gesteuert operieren sowie komplexe, unerwartete und teils nicht kalkulierbare Verhaltensweisen und Verfahren anwenden. Das bedeutet auch, dass diese möglicherweise die Entscheidungs- und Bekämpfungsautorität vom Menschen zunehmend auf autonome oder vollautonome Systeme übertragen. Die Gefährdung eigener Kräfte sowie die Dynamik und Letalität des Gefechtsfeldes werden dadurch zusätzlich erhöht.

204. Die Grundlagen für die Digitale Souveränität im GB BMVg gelten für KI. Dies schließt u. a. ein, KI-Systeme selbstständig und frei von ungewollter Einflussnahme Dritter betreiben und auf Basis eigener Daten weiterentwickeln zu können, auch in Ausübung staatlicher Hoheitsrechte. In dieser Hinsicht bezieht sich Souveränität auf den Grad der Kontrolle, den die Bw über die KI-Systeme, ihre KI-Infrastruktur, Datenverarbeitungsprozesse und die zugrundeliegenden Daten hat, die sie insbeson-

¹¹ Das BMVg definiert in Abstimmung mit dem Auswärtigen Amt (AA) LAWS wie folgt: Ein WaSys, welches in erster Linie dazu bestimmt ist, tödliche Gewalt allein gegen Personen zur Wirkung zu bringen, und welches, ohne jegliche menschliche Einflussnahme und Kontrolle sein Umfeld und seinen internen Zustand wahrnimmt, eine Beurteilung der Situation vornimmt, entscheidet, rational handelt, evaluiert und daraus lernt.

dere für militärische KI-Fähigkeiten entwickelt, einführt oder nutzt, einschließlich solcher, die (teil-) autonome Funktionen und Systeme ermöglichen. Dazu gehört eine angemessene Transparenz bei KI-Modellen, vor allem im Bereich der Entscheidungsunterstützung, zu gewährleisten sowie die hoheitliche Kontrolle und Sicherheit der im GB BMVg genutzten Daten herzustellen und dauerhaft zu bewahren. Synergieeffekte mit Institutionen der Inneren und Äußeren Sicherheit der Bundesregierung sind zu nutzen.

205. KI ist als Schlüsseltechnologie in der nationalen Strategie für die Sicherheits- und Verteidigungsindustrie festgelegt. In der Folge ist diese Festlegung zu operationalisieren und für nationale und Bündnisfähigkeiten entlang der Prozesse und Verantwortlichkeiten voranzutreiben.

206. Technologisch geschlossene, proprietäre Software-Lösungen sind dementsprechend möglichst zu vermeiden, um die Forderung nach angemessener Transparenz bei Algorithmen zu erfüllen sowie Kontrolle über Daten und die Handlungsfähigkeit im digitalen Umfeld aufrecht zu erhalten.

2.2 Rechtlicher Rahmen

207. Die Entwicklung, Einführung und Nutzung von KI im GB BMVg unterliegt dem anwendbaren nationalen, unionalen und internationalen Recht. Die rechtlichen Verpflichtungen sind in allen Bereichen einzuhalten, vom Arbeitsrecht, dem Personalvertretungsrecht, dem Datenschutz, dem Urheberrecht bis hin zum Humanitären Völkerrecht in bewaffneten Konflikten. Für das Handeln der Streitkräfte und der Bundeswehrverwaltung gelten bei der Anwendung von KI grundsätzlich die gleichen verfassungsrechtlichen Rahmenbedingungen, wie beim übrigen Handeln.

208. Die „Leitlinien für den Einsatz Künstlicher Intelligenz in der Bundesverwaltung“ finden Berücksichtigung, soweit die Besonderheiten des Verteidigungsressorts dies zulassen. Im Bereich der Verwaltungsabläufe existiert zudem mit § 9 EGovG eine gesetzliche Grundlage, wie bei der Einführung einer IT-Unterstützung oder einer wesentlichen Änderung derselben zu verfahren ist. Vor der Einführung von digitaler Unterstützung bzw. einer wesentlichen Änderung einer IT-Unterstützung soll eine Optimierung der Ablauforganisation bzw. der prozessualen Abläufe sicherstellen, dass entstehende organisatorische Optimierungspotenziale gehoben werden. Dabei können durch die Nutzung von KI Erkenntnissen zur Organisationsentwicklung gewonnen werden.

209. Die **EU KI-VO**¹² findet Anwendung, soweit keine der darin kodifizierten Bereichsausnahmen greift (siehe hierzu Anlage 5.1). Artikel 2 Absatz 3 Unterabsatz 2 EU KI-VO sieht vor, dass KI-Systeme,

¹² Die EU KI-VO trat am 1. August 2024 in Kraft, ist aber noch nicht vollumfänglich anwendbar. Am 19. November 2025 veröffentlichte die Europäische Kommission einen Vorschlag zur umfassenden Novellierung der Verordnung über Künstliche Intelligenz mit dem Ziel der Vereinfachung (Vorschlag zu einer Digital-Omnibus-Verordnung zur KI, EU-Kommission, COM (2025) 836 final). Das Verbot bestimmter KI-Systeme und die Anforderungen an die KI-Kompetenz begannen am 2. Februar 2025 zu gelten (Kapitel 1 und 2). Die Regeln für benannte Stellen (Kapitel 3, Abschnitt 4), GPAI-Modelle (Kapitel 5), Governance (Kapitel 7), Vertraulichkeit (Artikel 78), Sanktionen (Artikel 99 und 100) begannen am 2. August 2025 zu gelten. Mit Ausnahme von Artikel 6 Absatz 1 sollen die übrigen Bestimmungen am 2. August 2026 Anwendung finden. Artikel 6 Absatz 1 und die entsprechenden Verpflichtungen sollen am 2. August 2027 zu gelten beginnen.

wenn und soweit sie ausschließlich für militärische Zwecke, Verteidigungszwecke oder Zwecke der nationalen Sicherheit in Verkehr gebracht, in Betrieb genommen oder verwendet werden, vom Geltungsbereich der Verordnung ausgenommen sind (sogenannte Bereichsausnahme EU KI-VO).¹³ Dies ist u. a. notwendig, da z. B. mit Offenlegungspflichten Risiken entstehen, die mit dem Geheimschutz nicht vollständig vereinbar sind.

210. Um KI-Systeme entsprechend rechtssicher auszugestalten und im Projektverlauf Handlungssicherheit zu gewährleisten, ist es erforderlich, dass die Festlegung, ob ein KI-System in die Anwendung der Bereichsausnahme fällt (Betrachtung des Zwecks der beabsichtigten Nutzung des KI-Systems in einem Gesamtsystem), zum frühestmöglichen Zeitpunkt – im Regelfall in der Analysephase im Integrierten Projektteam (IPT) bzw. im Integrierten Clusterteam (ICT) – durch Zweckbindung getroffen wird. In der Fähigkeitslücke und Funktionalen Forderung (mit Lösungsvorschlag) bzw. im Clusterprogramm Teil 3 ist die Zweckbindung und die daraus resultierende Anwendung der entsprechenden Bereichsausnahme zu dokumentieren. Die Harmonisierung der Anwendung der Bereichsausnahmen mit den übrigen EU-Mitgliedstaaten ist anzustreben.

211. Gemäß Artikel 2 Absatz 6 gilt die EU KI-VO zudem nicht für KI-Systeme oder KI-Modelle, einschließlich ihrer Ausgabe, die eigens für den alleinigen Zweck der wissenschaftlichen Forschung und Entwicklung entwickelt und in Betrieb genommen werden.

212. Die EU KI-VO nimmt mit Bezug auf KI eine Risikoeinteilung vor, unterteilt diese in verbotene Systeme (Artikel 5 EU KI-VO) und Hochrisiko-Systeme (Artikel 6 EU KI-VO) sowie Systeme mit normalem und minimalem Risiko. Je nach Risikostufe sieht die VO einen breiten Pflichtenkatalog vor. Diese Pflichten treffen primär Hersteller/Anbieter/Einführer, teilweise aber auch die Betreiber der KI. Für den Auftrag der Streitkräfte sind unter Geheimhaltungsaspekten insbesondere Offenlegungspflichten problematisch. Jedoch stellen auch Dokumentations- und Transparenzpflichten nicht zu unterschätzende bürokratische Hürden auf, welche den erforderlichen schnellen Aufwuchs an Fähigkeiten verzögern können. Es gilt daher für die Nutzung von KI in der Bw stets zu prüfen, inwieweit die in Artikel 2 (3) EU KI-VO normierte Bereichsausnahme anwendbar ist (siehe Anlage 5.1).

¹³ Eine Stellungnahme zur inhaltlichen Differenzierung der Bereichsausnahme EU KI-VO ist der Anlage 5.1.1 zu entnehmen.

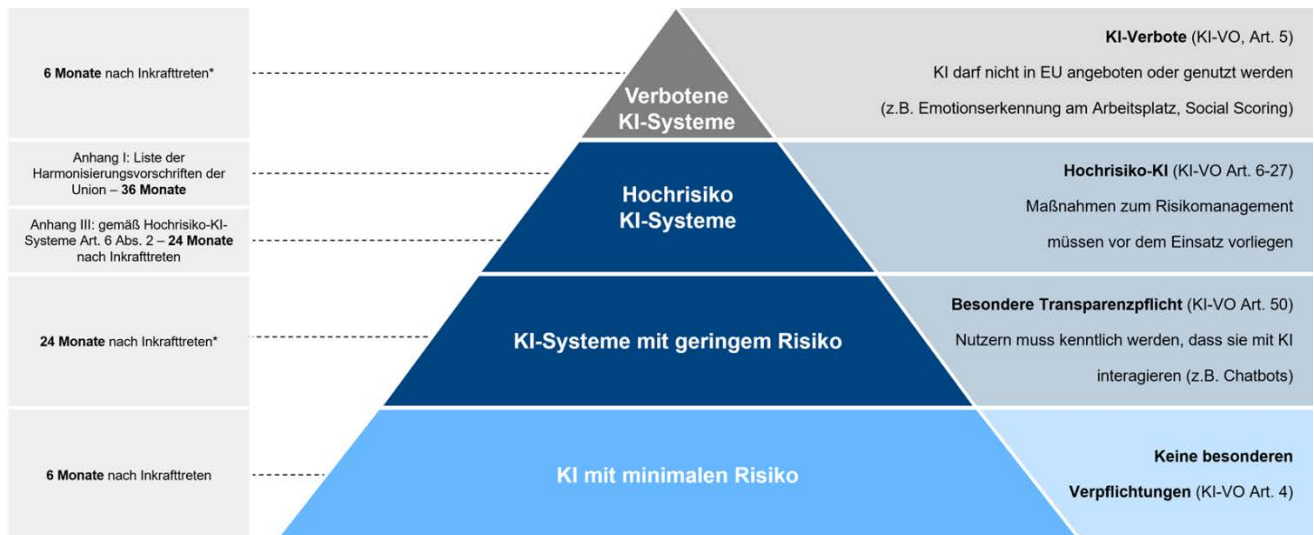


Abb. 1: Anwendung EU KI-VO

213. Besondere Bedeutung kommt der **Sicherheit und Zuverlässigkeit von KI-Systemen** zu. Sie müssen robust gestaltet, regelmäßig automatisiert getestet und validiert werden, um potenzielle Risiken bei der Anwendung von KI zu minimieren. Eine laufende, automatisierte Risikoüberwachung ermöglicht, auf mögliche Schwachstellen und unvorhergesehene Auswirkungen frühzeitig zu reagieren. Der Prozess der automatisierten Risikoüberwachung, Testung und Validierung (inklusive ggf. erforderlichlichem Nachtraining) muss einmalig grundsätzlich entworfen und dann von Projekten mit KI-Anteilen zur Zertifizierung vorgelegt werden.

214. Der **Schutz sensibler Daten** wird durch technische und organisatorische Maßnahmen sichergestellt, um den Anforderungen an Vertraulichkeit und Integrität gerecht zu werden.

215. Die **Datenschutz-Grundverordnung (DSGVO)**^{14, 15} findet Anwendung; Datenschutz muss für die Bw – auch bei der Entwicklung, Einführung und Nutzung von KI – entsprechend beachtet werden.

216. Sozialdiskriminierende Entscheidungen von KI-Systemen müssen unterbunden und Algorithmen regelmäßig automatisiert auf Verzerrungen überprüft werden.¹⁶

¹⁴ Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung). Sie findet trotz Artikel 2 Absatz 2 lit. b DSGVO über § 1 Absatz 8 Bundesdatenschutzgesetz (BDSG) Anwendung.

¹⁵ Siehe auch: KI in Behörden – Datenschutz von Anfang an mitdenken (Handreichung der BfDI für die Bundesverwaltung, Stand: 22.12.2025).

¹⁶ Siehe auch das „Gutachten der Datenethikkommission der Bundesregierung“ (DEK), Berlin, Oktober 2019. Aus Sicht der DEK ist KI lediglich eine besondere Ausprägung algorithmischer Systeme und teilt viele ethisch und rechtlich relevante Eigenschaften mit anderen Arten solcher Systeme, weshalb die DEK ihre Ausführungen auf Daten und algorithmische Systeme allgemein bezieht.

217. Die Einführung eines WaSys mit KI-Anteilen in die Bw erfolgt entlang der Anforderungen des Artikels 36 des Zusatzprotokolls I von 1977 zu den Genfer Abkommen von 1949, die im Rahmen der vom BMVg veranlassten Waffenprüfung sicherzustellen ist.

218. Mögliche Kontrollsysteme, die sicherstellen sollen, dass die Nutzung von KI-Systemen im Einklang mit nationalen, unionalen und internationalen rechtlichen Verpflichtungen steht, unterstützen das für die KI-Nutzung verantwortliche Personal. Solche Kontrollen können zu jedem Zeitpunkt im Entwicklungs- und Nutzungszyklus des KI-Systems erfolgen und können auch die Verwendung personenbezogener Daten umfassen, wenn das KI-System in einem internationalen Kontext eingesetzt wird, wo gesetzliche Verpflichtungen von denen DEU abweichen oder mit diesen in Konflikt geraten könnten.

219. Dem Ansatz der Clusterlogik folgend können wiederverwendbare KI-Systeme den notwendigen Prüfaufwand reduzieren. Hierzu sind fachseitig geeignete Nutzungsbereiche für KI so zu identifizieren, dass eine Wiederverwendbarkeit der Modelle ermöglicht wird.

2.3 Ethischer Rahmen

220. Die rasante Entwicklung von KI birgt erhebliche Potenziale für militärische Operationen und erfordert damit auch eine ethische Reflexion.¹⁷ Die Nutzung von KI erfordert eine sorgfältige Berücksichtigung ethischer Grundsätze, muss mit den Werten, Normen und Prinzipien der Inneren Führung im Einklang stehen und setzt eine sorgfältige Abwägung technologischer Möglichkeiten und militärischer Verantwortung voraus. Normative Präferenzänderungen in Krisen- und Konfliktzeiten sind dabei zu berücksichtigen. Eine kontinuierliche Bewertung und ggf. Anpassung ethischer Prinzipien in Abstimmung mit völkerrechtlichen und internationalen Rahmenbedingungen ist erforderlich, um mit der technologischen Entwicklung Schritt zu halten und einem Gegner gegenüberzutreten zu können, dessen (Kampf-) Handlungen sich nicht mit unseren Werten vereinbaren lassen.

221. Governance und NATO-Prinzipien: Die Implementierung einer KI-Governance (siehe Abschnitt 4.1) stellt sicher, dass ethische Prinzipien und rechtliche Vorgaben in den gesamten Lebenszyklus von KI-Systemen integriert werden. Die Bw orientiert sich an den „**NATO Principles of Responsible Use of Artificial Intelligence in Defence**“ (PRU), die einer Governance folgend Rechtmäßigkeit, Verantwortlichkeit und Rechenschaftspflicht sowie Nachvollziehbarkeit (Erklärbarkeit und Rückverfolgbarkeit), Zuverlässigkeit und Vorurteilsminderung fordern. Damit kein Verantwortungsbruch erfolgt, wenn Dritte KI-Systeme für die Bw entwickeln und entsprechende Designentscheidungen treffen, muss sichergestellt werden, dass beschaffte KI-Systeme auf ethische Prinzipienwahrung entsprechend den NATO-Vorgaben hin validiert sind.

¹⁷ Die Universität der Bundeswehr Hamburg, Lehrstuhl für Technikethik, interdisziplinärer Forschungshub KI-Ethik (AI Ethics Research Hub), widmet sich der in diesem Unterkapitel dargestellten ethischen Dimension von KI, siehe <https://www.hsu-hh.de/iai/> und <https://www.hsu-hh.de/theevs/ai-ethics> abgerufen am 27. August 2025.

222. Um die Einhaltung der PRU zu gewährleisten beabsichtigt der GB BMVg das „NATO Responsible AI (RAI) Assessment and Toolkit“ mit dem Ziel eines einheitlichen Vorgehens im Bündnis anzuwenden, sofern die derzeit laufenden Tests der NATO positiv abgeschlossen werden. Diese Aktivitäten sind durch den GB BMVg aktiv zu begleiten und Erkenntnisse frühzeitig in laufenden Projekten zur Anwendung zu bringen, um die Interoperabilität mit den Bündnispartnern gemeinsam zu gestalten.

223. Vertrauenswürdigkeit: Im Zusammenhang mit KI bezieht sich Vertrauenswürdigkeit spezifisch auf Fairness, Autonomie und Kontrolle, Transparenz, Verlässlichkeit, Sicherheit und Datenschutz. Die jeweiligen Ausprägungen der spezifischen Kriterien hängen hierbei stark vom Einsatzgebiet ab und müssen auf individuelle Szenarien angepasst werden, um die Qualität und Effizienz des KI-Systems zu maximieren, während gesellschaftliche Risiken und unerwünschte Nebeneffekte minimiert werden sollen. Während für militärische KI-gestützte Entscheidungsprozesse die menschliche Verantwortung betont wird, gelten für administrative Anwendungen zusätzliche, darunter auch soziokulturelle Kriterien, wie bspw. Diskriminierungsfreiheit. In beiden Fällen ist die Nachvollziehbarkeit der Entscheidungsprozesse ein entscheidendes Kriterium, um die Vertrauenswürdigkeit der Systeme sicherzustellen.

224. Menschenzentrierter Ansatz, Verantwortlichkeit und menschliche Kontrolle: Ein zentrales Prinzip ist der menschenzentrierte Ansatz, bei dem der Mensch die Kontrolle über kritische Entscheidungen¹⁸ behält. Die KI dient hiernach als unterstützendes Werkzeug bei der Entscheidungsfindung, nicht selbst als autonomer Entscheider. Die Verantwortung für Entwicklung, Nutzung und Ergebnisse von KI-Systemen muss klar definiert sein, um Transparenz, Vertrauen, Zurechenbarkeit und Sicherheit zu gewährleisten. Dazu gehört auch die Offenlegung von KI-Ergebnissen und der zur Entwicklung genutzten Datengrundlage durch angemessene Dokumentation der Verhaltensweisen technischer Systeme und deren Möglichkeit zur Überprüfung. KI-Systeme sind ausreichend zu testen. Dazu sind akzeptierte Fehlertoleranzen festzulegen. Die revisionssichere Speicherung aller Aktivitäten und Anpassungen in der Nutzung des Systems kann die Rückverfolgbarkeit und die Nachvollziehbarkeit der Ergebnisse gewährleisten, sofern dies geboten ist. Regelmäßige Audits stellen sicher, dass ein KI-System auch im laufenden Betrieb den ethischen, rechtlichen und technischen Standards entspricht. Personal muss über die Kompetenz verfügen, eine Plausibilitätsprüfung der KI durchzuführen zu können, um fehlerhafte Entscheidungsvorschläge zu identifizieren und entsprechend zu handeln.

225. Durch die ständige sowie falsche Nutzung von KI können Fähigkeiten beim Personal verloren gehen, die bei einem Ausfall von KI dann nicht mehr kompensiert werden können. Daher ist in kritischen Bereichen zwingend ein In-Übung-halten ohne die Nutzung von KI notwendig, um eine nötige Resilienz gegen beispielsweise Störungen durch den Gegner aufrecht zu erhalten. Im StVwH ist die Nutzung von KI-Sprachmodellen nicht als reiner „Effizienzverstärker“ vorzusehen, sondern im Schwerpunkt als

¹⁸ Dies gilt nicht nur für KI auf dem Gefechtsfeld, sondern insbesondere auch für KI-Systeme im Personalmanagement.

Instrument zur Qualitätssicherung und Reflexion. Es müssen immer die Schritte der eigenen Urteilsbildung sichtbar bleiben, um Kompetenzabbau vorzubeugen.

226. DEU hat sich politisch verpflichtet¹⁹, dass die Bw sicherstellen muss, die Entwicklung, Einführung und Nutzung von KI-Systemen unter den gebotenen Vorkehrungen in Bezug auf kontextangemessenes menschliches Urteilsvermögen und kontextangemessene Kontrolle über die Anwendung militärischer Gewalt sowie im Rahmen einer verantwortlichen Befehlskette sicherzustellen (siehe Ausführungen im Abschnitt 3.3.5).

227. Menschliche Kontrolle ist dabei nicht gleichzusetzen mit menschlicher Interaktion. Wenngleich menschliche Interaktion bei Zielauswahl und -bekämpfung, wann immer möglich, stattfindet, erfordert nicht jeder Wirkmitteleinsatz zwingend eine menschliche Interaktion. Der Grad menschlicher Interaktion ist vom konkreten Operationskontext abhängig. In Luftkampfsituationen oder auf See, also fernab von zivilen Personen und Infrastrukturen, kann das notwendige Maß an menschlicher Interaktion ggf. niedriger sein als beim Einsatz etwa in urbanen Gebieten.

228. Nachhaltigkeit und Zertifizierung: Für KI-Systeme, die in den Anwendungsbereich der EU KI-VO und der KI-Leitlinien für die Bundesverwaltung fallen, bevorzugt der GB BMVg künftig²⁰ KI-Systeme, die entsprechend der EU KI-VO eine CE-Zertifizierung aufweisen. Bei Anwendung der Bereichsausnahme der EU KI-VO beschafft und nutzt der GB BMVg ausschließlich KI-Systeme, die rechtlichen sowie ethischen Standards (siehe Nr. 221) entsprechen.

2.4 Internationaler Rahmen

229. Angesichts der globalen KI-Entwicklungen²¹ ist die internationale Zusammenarbeit von entscheidender Bedeutung; dieses KD dient zugleich zur Unterstützung der Positionierung auf internationaler Ebene. Die Abstimmung mit internationalen Partnern und die Einhaltung globaler Normen und Standards tragen dazu bei, eine verantwortungsvolle, ethische und rechtmäßige Nutzung von KI im militärischen Kontext sicherzustellen. Durch die – teils führende – Mitarbeit in internationalen Initiativen zur verantwortungsvollen Nutzung von KI im Verteidigungsbereich gestaltet DEU die Formulierung und Implementierung entsprechender Normen und Standards mit.²² Dies fördert nicht nur das Vertrauen zwischen Nationen, sondern hilft auch dabei, gemeinsame Normen und Standards zu etablieren und

¹⁹ Paris Declaration on Maintaining Human Control in AI-enabled Weapon Systems. Élysée. (<https://www.elysee.fr/emmanuel-macron/2025/02/11/paris-declaration-on-maintaining-human-control-in-ai-enabled-weapon-systems> (abgerufen am 17. August 2025))

²⁰ Zertifizierungen gemäß EU KI-VO erfolgen ab spätestens August 2026.

²¹ Das **Defense AI Observatory** an der Universität Bw Hamburg beobachtet und analysiert den internationalen Einsatz von KI durch Streitkräfte und stellt entsprechende Produkte unter <https://defenseai.eu/> zur Verfügung, abgerufen am 17. August 2025.

²² Hervorzuheben sind hier die Mitwirkungen bei den Bezügen Political Declaration on Responsible Military Use of Artificial Intelligence and Autonomy, Responsible AI in the Military Domain (REAIM), United Nations Institute for Disarmament Research (UNIDIR) Expert Network on Governance of Artificial Intelligence in the Military Domain, VN Resolutionen A/Res/79/239 (2024) und A/Res/80/58 (2025) und Paris Declaration on Maintaining Human Control in AI-enabled Weapon Systems.

durchzusetzen. Gleichzeitig soll DEU KI-Souveränität gestärkt und die glaubwürdige Abschreckung durch überzeugende Kampf- und Siegfähigkeit gegenüber möglichen NATO-Gegnern sichergestellt werden.

230. DEU bringt sich in die Erarbeitung nationaler und internationaler Normen und Standards ein und implementiert diese im Rahmen des Planungs-, Rüstungs- und Nutzungsprozesses. Ziel ist die Gestaltung eines Normenrahmens in der NATO und EU, der zu einem Verifizierungs- und Validierungsprozess führt, der Interoperabilität und Normerfüllung im Verständnis der westlichen Wertegemeinschaft nachhält. Nationales Risikomanagement (RM), Testung, Evaluierung und Validierung orientieren sich grundsätzlich an diesem angestrebten Normenrahmen. Nationale Erweiterungen werden nur implementiert, wenn diese aus nationaler Sicht rechtlich oder politisch unumgänglich sind, um Erfüllungsaufwand zu minimieren und die Interoperabilität zwischen nationalen KI-Systemen bereits in der Entwurfsphase als designbestimmendes Merkmal zu implementieren. DEU unterstützt zudem in den VN die Entwicklung eines Normenrahmens, um gemeinsame Standards innerhalb der globalen Wertegemeinschaft zu etablieren.²³

²³ Siehe hierzu die Resolutionen der VN-Generalversammlung A/RES/78/265 (2024), Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development in Verbindung mit A/RES/80/58 (2025), Artificial intelligence in the military domain and its implications for international peace and security und das von UNIDIR Anfang 2026 gegründete Framework of Responsible Industry Behaviour for AI in the Military Domain.

3 Handlungsfelder

301. Die zuvor beschriebenen Rahmenbedingungen bilden grundsätzlich die Leitplanken für die nachfolgend beschriebenen Handlungsfelder.

3.1 Handlungsfeld Ausbildung und Qualifizierung

302. Ausbildung und Qualifizierung zum Erwerb einer erforderlichen KI-Kompetenz über den gesamten Lebenszyklus von in die Bw einzuführender oder eingeführter IT oder sonstiger (Waffen-) Systeme ist im Rahmen der etablierten Ausbildungs- bzw. Qualifizierungslandschaft und -verfahren im GB BMVg abzubilden. Den Vorgaben und Handlungsempfehlungen der „Leitlinien für den Einsatz Künstlicher Intelligenz in der Bundesverwaltung“, der NATO KI Strategie und, soweit anwendbar, der EU KI-VO²⁴ wird gefolgt.

303. Die Grundsätze und strategischen Ziele für die Ausgestaltung der „Ausbildung und Qualifizierung im Bereich der IT im GB BMVg (AuQITGBBMVg)“ sind durch die bzw. den Ressort-Chief Information Officer (CIO) geregelt. Diese konkretisieren konzeptionelle Vorgaben für die militärfachliche Ausbildung der Soldatinnen und Soldaten sowie für die Qualifizierung der zivilen Beschäftigten der Bw aus IT-fachlicher Sicht.

304. . Im GB BMVg bereits etablierte Ausbildungs- und Qualifizierungsmaßnahmen zu KI sind diesbezüglich zu überprüfen und ggf. anzupassen.

305. Es werden für die Ausbildung und Qualifizierung von Angehörigen des GB BMVg vier Zielgruppen unterschieden:

- Allgemeine Ausbildung und Qualifizierung zum grundsätzlichen KI-Kompetenzerwerb für alle Angehörigen des GB BMVg,
- erweiterte Ausbildung und Qualifizierung für DP-unabhängigen Kompetenzerwerb,
- erweiterte Ausbildung und Qualifizierung für Führungskräfte des GB BMVg und
- dienstposten-/dienstpostentypspezifische Ausbildung und Qualifizierung, die der Wahrnehmung von Aufgaben und Rollen unter Nutzung von KI²⁵ dient.

3.1.1 Allgemeine Ausbildung und Qualifizierung

306. Aufgrund der immer weiter zunehmenden Bedeutung von KI sowohl im öffentlichen Umfeld als auch im GB BMVg ist es erforderlich, allen Angehörigen im GB BMVg mit Zugang zum IT-System der Bw dienstposten- und aufgabenunabhängig eine allgemeine Basiskompetenz über KI zu vermitteln. Hierbei sind neben einem allgemeinen Überblick über die Chancen, Potenziale, Möglichkeiten bis hin zu anwendbaren Arbeitserleichterungen von KI, die rechtlichen und ethischen Grundlagen zu berück-

²⁴ EU KI-VO, Artikel 2 Absatz 3 in Verbindung mit Artikel 4.

²⁵ Einschließlich Forschung, Entwicklung, Erprobung.

sichtigen. Das Basistraining soll eine Grundsensibilisierung schaffen, indem es den bewussten persönlichen Umgang mit KI im GB BMVg fördert. Es umfasst die durch übergeordnete Vorgaben (siehe Abschnitt 3.2.1) festgelegten Inhalte:

- Definition und Funktion von KI-Systemen,
- Potenziale und Risiken von KI-Nutzung im GB BMVg,
- Ethische und rechtliche Grundsätze im Umgang mit KI im GB BMVg und
- Hinweise auf geltende interne Regelwerke und Informationsplattformen zum Einsatz von KI im GB BMVg.

307. Die Inhalte des Basistrainings für alle Angehörigen des GB BMVg sind aufwandsarm umzusetzen. Möglichkeiten bieten hierzu Web-Based Trainings, In-House-Schulungen oder E-Learning Module. Zeitintensive Trainingsmaßnahmen sind zu vermeiden.

3.1.2 Erweiterte Ausbildung und Qualifizierung

308. Angehörige des GB BMVg, welche wiederkehrend, aktiv und gezielt KI in ihrem dienstlichen Alltag einsetzen, deren Ergebnisse in Entscheidungsfindungsprozesse einfließen lassen sowie deren Einsatz steuern und bewerten, sind aufbauend auf dem Basistraining nach dem Abschnitt 3.1.1 erweitert zu qualifizieren. Ziel ist die Sensibilisierung, sachgerechte Einordnung von Informationen und allgemeine Handlungssicherheit im Umgang mit dienstlich bereitgestellten Anwendungen und KI-Systemen.

309. Der Einsatz dieser KI-Systeme ist nicht abhängig vom Dienstposten oder der Aufgaben am Dienstposten. Das erweiterte Training enthält mindestens folgende Inhalte²⁶:

1. Grundlagen der KI

- Methoden des maschinellen Lernens und Bestandteile von KI-Technologien,
- Anwendungsbereiche, Service/Software-Angebot und Mehrwert in der Praxis,
- Leistungsvermögen und Umgang mit den Grenzen von KI-Technologien und
- Grundlagen der Nutzung (inklusive Möglichkeiten des Promptings).

2. Rechtliche Rahmenbedingungen²⁷

- EU KI-VO (inklusive Bereichsausnahmen) und nationale Umsetzung,
- DSGVO und BDSG im Kontext von KI,
- Urheberrecht im Kontext KI,
- Arbeitsrecht und Personalvertretungsrecht im Kontext KI,
- Rechtsstaatsprinzip mit Nachvollziehbarkeit des Verwaltungshandelns,

²⁶ Die Inhalte sind in Abhängigkeit von den technischen Entwicklungen bei KI-Systemen und der Standardsetzung bei EU und NATO bedarfsweise anzupassen.

²⁷ Siehe Abschnitt 2.2.

- Haftungsfragen bei Fehlern durch KI-Systeme und
- KI und Humanitäres Völkerrecht.

3. Ethische und gesellschaftliche Aspekte²⁸

- Voreingenommenheit (Bias) und Diskriminierung in KI-Systemen²⁹,
- Transparenz und Nachvollziehbarkeit,
- Menschliche Autonomie und Kontrolle und
- Verantwortung bei der Nutzung von KI (Do's and Don'ts).

4. Technische Kompetenz

- Grundlegendes Verständnis von Algorithmen und Datenverarbeitung für Angehörige des GB BMVg ohne einschlägige technische KI-Vorkenntnisse,
- Kenntnis von und Umgang mit KI-Tools und KI-Infrastruktur im GB BMVg und
- Anwendung von Methoden des maschinellen Lernens und der Bereitstellung von KI-Software-artefakten.

5. Risikomanagement³⁰

- Identifikation und Bewertung von Risiken,
- Maßnahmen zur Minimierung von Fehlern und Missbrauch und
- Monitoring und regelmäßige Überprüfungsmöglichkeiten von KI-Systemen.

3.1.3 Ausbildung und Qualifizierung von Führungskräften

310. Um alle Angehörigen des GB BMVg angemessen und zukunftsorientiert bei dem Einsatz und der Nutzung von KI begleiten und führen zu können, benötigen Führungskräfte ein fundiertes Verständnis von KI und ihrer Wirkungsweise auf die Organisation, die Prozesse und Abläufe.

311. Im Rahmen der Ausbildung und Qualifizierung von Führungskräften ist ein besonderer Fokus zu legen auf

- Kompetenzen zum strategischen und steuernden Einsatz von KI,
- rechtliche und ethische³¹ Rahmenbedingungen, um den Einsatz von KI freigeben, stoppen oder ausweiten zu können,
- Verantwortlichkeiten der Führungskraft und die Vorbildfunktion, um den realen KI-Umgang im Dienstalltag zu prägen und

²⁸ Siehe Abschnitt 2.3.

²⁹ Berücksichtigung United Nations Security Council Resolution (UNSCR) 1325 ff.

³⁰ Siehe Abschnitt 3.3 und Anlage 5.2.

³¹ Die ethischen Implikationen, die sich aus der Nutzung von KI ergeben, sind vor allem vor dem Hintergrund der aktuellen Herausforderungen der Bw in der Landes- und Bündnisverteidigung (LV/BV) dabei besonders zu berücksichtigen.

- Metakompetenzen zu: Konfliktfähigkeit (Durchsetzen des Einsatzes/des Nicht-Einsatzes von KI-Systemen) und langfristige Folgenabschätzung des Einsatzes von KI (Kompetenzverlust infolge Automatisierung usw.).

312. Im Gegensatz zu den in den Abschnitten 3.1.1 und 3.1.2 dargelegten Ausbildungsforderungen, welche die Absicht verfolgen die korrekte Nutzung von KI zu vermitteln, soll die Ausbildungsforderung für Führungskräfte, diese befähigen KI zu erlauben, zu begrenzen – und dafür einzustehen.

3.1.4 Aufgabenbezogene Ausbildung und Qualifizierung

313. Der über die allgemeine Basiskompetenz hinausgehende Kompetenzerwerb und -erhalt ist essenziell für die Aufgabenwahrnehmung der jeweiligen Bundeswehrangehörigen in den Prozessen des GB BMVg.

314. Angehörige des GB BMVg, welche dienstposten- und/oder aufgabenbezogene KI-Systeme anwenden, müssen die Ergebnisse kritisch bewerten und interpretieren können.

315. Für die Integration von KI-Systemen im dienstlichen Alltag sind die erforderlichen Aufgaben zu identifizieren und mit Rollen dienstpostenbezogen zu hinterlegen. Im Rahmen der identifizierten Aufgaben sind die Bedarfs- und Mengengerüste festzustellen und die Ausbildungs- bzw. Qualifizierungsmaßnahmen im Rahmen der bestehenden Verfahren zu implementieren. Dabei ist zu prüfen, ob neue Ausbildungs- und Qualifizierungsmaßnahmen geplant werden oder ob auf Bestehendem aufgesetzt wird.

316. Dies beinhaltet grundsätzlich auch die (waffen-)systemspezifische Ausbildung. Diese ist im Rahmen eines projektbezogenen Ausbildungskonzepts zu bearbeiten.

3.1.5 Künstliche Intelligenz in der Ausbildung und Qualifizierung

317. Ziel der Integration von KI in die Virtuelle Lernumgebung der Bundeswehr (VLBW) ist die Verbesserung der Wissensvermittlung und die Unterstützung der Ausbildenden, Lernenden, Ausbildungsmanagerinnen und Ausbildungsmanagern sowie Entscheidungstragenden durch Automatisierung und Personalisierung. Dies kann insbesondere erreicht werden durch

- individuelle Bereitstellung von Ausbildungs- bzw. Qualifizierungsinhalten, angepasst an den Kenntnisstand und das Lerntempo der Nutzenden (mögliche Verkürzungen der individuellen Trainingsdauer können dabei in der Ausbildung/Lehrgangsplanung mitberücksichtigt werden),
- Optimierung der Lernprozesse, um eine effizientere Nutzung der Lernzeit und eine höhere Motivation der Nutzenden zu erreichen und
- Unterstützung der Funktionstragenden in der Ausbildung und Qualifizierung durch KI-gestützte Werkzeuge, die u. a. repetitive Aufgaben reduzieren, gezielte Interventionen sowie präzises Feedback ermöglichen und die Erzeugung von digitalen Inhalten unterstützen.

318. Die Möglichkeiten der Implementierung des adaptiven Lernens mittels KI kann die Brücke zwischen dem individuellen Lernen, dem kompetenzorientierten Ansatz in der Ausbildung und der Unterstützung des Personalmanagements darstellen.

319. Die Implementierung von KI-Systemen, welche den vorangestellten Zielen und Absichten dienen, ist voranzutreiben.

3.2 Handlungsfeld Fähigkeiten

320. KI-Systeme leisten einen Beitrag zur Einsatzbereitschaft und zur Zukunftsfähigkeit des GB BMVg. Für die Fähigkeitsentwicklung bedeutet dies, dass die Nutzung von KI-Systemen zu neuen Fähigkeiten führen, vorhandene Fähigkeiten ergänzen, aber auch bisherige Fähigkeiten überflüssig machen kann (siehe Abschnitt 1). Dabei müssen KI-Systeme in und für alle Dimensionen und über alle Planungskategorien berücksichtigt werden.

321. Ein künftig querschnittlich höherer Grad an KI-Unterstützung ist unerlässlich, um mit gegnerischen Kräften Schritt zu halten bzw. ihnen voraus zu sein, eigene personelle Begrenzungen zu kompensieren und die Wirkung und den Schutz des eigenen Personals, der eigenen Einrichtungen und von eigenen Systemen zu erhöhen.

3.2.1 Vorgaben für die Fähigkeitsentwicklung der Bereiche

322. KI-Unterstützung ermöglicht moderne Arbeitsmethoden und entlastet Personal. Dies kann zu mehr Arbeitszufriedenheit und einer verbesserten Personalgewinnung und Personalbindung führen. Zudem erlaubt KI eine Konzentration auf höherwertige Tätigkeiten und bietet dadurch persönliche Entwicklungsmöglichkeiten. Die organisatorischen Auswirkungen auf Aufgaben und Rollen ist vor der Einführung zu prüfen. Die Implementierung von KI-Systemen in Geschäftsprozessen erfordert regelmäßig die Einführung eines ablauf- und aufbauorganisatorisch optimierten Sollprozesses. Auswirkungen von KI auf die Arbeitszufriedenheit und mögliche soziale gesundheitliche Auswirkungen sind zu beobachten und mit zeitlichem Abstand zur Einführung von KI-Systemen in Form ganzheitlicher Studien wissenschaftlich zu evaluieren und zu steuern.

323. Die KI wird wesentliche Aspekte künftiger Einsatzführung verändern, indem die Funktionskette vom Sensor zum Effektor optimiert wird. So gestaltet sich das Gefecht präziser und schneller, aber auch weiträumiger und variabler. Optimierungen hinsichtlich Automatisierung und Datenübertragung sowie -verarbeitung mittels der Nutzung von KI sowie der Einsatz völlig neuer Wirkkonzepte in Mischformen (hybride Teams aus Menschen und KI im Manned-Unmanned Teaming) sind kontinuierlich zu beobachten, zu evaluieren und anlassbezogen einzuleiten. Bei der Einführung von KI ist hierbei der teilweise oder komplette Ausfall der KI-Unterstützung oder erforderlicher Netzwerke mitzudenken und entsprechende Alternativverfahren zu entwickeln und zu schulen.

324. Die Auslieferung von KI-System-Softwarebestandteilen, deren (Weiter-)Entwicklung (insbesondere maschinelles Lernen) und Betrieb (die Nutzung) sind dabei über Qualitätssicherungsmaßnahmen, Freigabeprozesse und technisch durch getrennte Plattformen der KI-Infrastruktur (Foundation) grundsätzlich voneinander zu trennen, um nicht definierte Betriebszustände zu vermeiden, Kontrolle aufrechtzuhalten und das Vertrauen in die Technologie und KI-Systeme zu stärken. Dies ist bei der Konzeption, Entwicklung und Beschaffung von KI-Systemen und der KI-Infrastruktur in der Bw zu berücksichtigen.

325. Bei der Beschaffung von Foundation Models sind die Kriterien der Wirtschaftlichkeit, der schnellen Verfügbarkeit und des sicheren Betriebs zu beachten. Verfügbare (vortrainierte) Modelle müssen in einer eigenen Experimental-Umgebung fortlaufend erprobt und bei Bedarf mit spezifischen Daten des GB BMVg für vielfältige eigene Einsatzbereiche adaptiert werden. Bei der Auswahl dieser Foundation Models sind die Vorgaben der Bundesregierung zur Stärkung der Sicherheits- und Verteidigungsindustrie, zur Nutzung von offenen Standards und vorrangiger Anschaffung von Open-Source-Software³² sowie für die Fähigkeitsentwicklung im Bereich von Schlüsseltechnologien und der Digitalen Souveränität zu berücksichtigen.

326. Das Erschließen des vielfältigen Anwendungspotenzials von KI-Systemen sowohl für die Weiterentwicklung der Fähigkeiten der Teilstreitkräfte (TSK) wie auch die Sicherstellung der bedarfsgerechten und durchhaltefähigen Auftragserfüllung der unterstützenden Organisationsbereiche (OrgBer), Kommandos und Dienststellen (DSt) steht erst am Anfang. Dies wird sich mit der technologischen Entwicklung stetig ausweiten und muss deshalb in allen Bereichen weiterhin möglichst umfassend und fortlaufend analysiert und erschlossen werden. Es gilt, die einschlägige Forschung, Entwicklung und Anwendung der weltweit beteiligten Institutionen und Akteure breit und hinreichend tief zu verfolgen, um Potenziale, Möglichkeiten, Grenzen und Risiken für die Bw zu erkennen.

327. Die OrgBer und die dem BMVg unmittelbar unterstellten Ämter und DSt sind gefordert, sich mit der Nutzung von KI zur Verbesserung seiner Fähigkeiten bzw. des Unterstützungsauftrages auseinanderzusetzen.

328. Die TSK Cyber- und Informationsraum (CIR) koordiniert entsprechende KI-Nutzungs- und Weiterentwicklungsansätze, identifiziert Synergieeffekte und gewährleistet die dimensionsübergreifende Entwicklung und Zusammenarbeit sowie die Anschlussfähigkeit der zivilen OrgBer.

329. Die Integration von KI in WaSys und das Herstellen einer Updatefähigkeit implementierter Algorithmen, Modelle und Daten sind gemäß des konzeptionellen Rahmens von SDD sicherzustellen.

330. Zum Erreichen einer digitalen Souveränität ist Open Source ein Baustein.

331. Multinationale Fähigkeitsentwicklung, Interoperabilitäts- und Standardisierungsaspekte in NATO, EU sowie bei den Partnernationen (u. a. Vereinigte Staaten von Amerika (USA), Frankreich

³² EGovG § 16a Open Source.

(FRA), Vereinigtes Königreich (GBR)) sind umfassend zu berücksichtigen, um die Interoperabilität der Bw als Streitkräfte im Kontext internationaler Einsätze sowie bei LV/BV weiter gewährleisten zu können. Dazu sind geeignete Schnittstellen zu implementieren. Das Framework Federated Mission Networking (FMN) kann den geeigneten Rahmen bilden.

332. Im Rahmen des Innovationsmanagements sind Entwicklungen im Bereich KI sowie deren Interdependenz zum Quantencomputing zu betrachten, insbesondere in Bezug auf Krypto- und Informationssicherheit.

3.2.2 Steuerung, Planung Teilportfolio Cyber/IT und Beschaffung

333. Kurze Entwicklungszyklen im Bereich KI werden durch Nutzung von Digitalisierungsplattform und Clusterlogik ermöglicht. KI-Fähigkeiten werden im Wirkverbund der Digitalisierungsplattform GB BMVg gesteuert, realisiert und betrieben, um standardisierte und wiederverwendbare KI-Services bereitzustellen. Involvierte DSt müssen fachlich und organisatorisch für die Zusammenarbeit mit zentralen Aufgabenträgern aufgestellt sein (siehe Abb. 2).

334. Um Synergien zu nutzen, Redundanzen zu vermeiden und eine medienbruchfreie Nutzung von Daten zu ermöglichen, soll die Finanzierung von querschnittlich Cluster- oder OrgBer-übergreifend genutzten KI-Systemen und -Services weitgehend über das Clusterprogramm³³ Analytics & Simulation (A&S) erfolgen. Die Regeneration von Experimentalplattform, sowie Test-, Entwicklungs- und Produktionsumgebung des A&S Service Mesh muss kontinuierlich über die Digitalisierungsplattform erfolgen. Dies erfolgt häufiger durch die Bereitstellung neuer Software und in größeren Abständen durch die Bereitstellung neuer Hardware u. a. als Teil der private Cloud Bw (pCloudBw).

335. In Projekten mit KI-Anteil in der Analyse- oder Realisierungsphase sind, wenn wirtschaftlich sinnvoll, die Vorgaben der Digitalisierungsplattform auch für KI-Systeme obligatorisch vertraglich festzuschreiben und die Vorgaben aus der Datenstrategie des GB BMVg zu berücksichtigen. Dies ist erforderlich, um das Potenzial der Digitalisierung und der Datenstrategie gerade auch für KI-Systeme zu nutzen.

3.2.3 KI-Infrastruktur

336. Der GB BMVg benötigt für einen harmonisierten und betreibbaren Aufwuchs von Fähigkeiten mit KI-Anteilen eine horizontal durchgängige KI-Infrastruktur, welche unter Federführung (FF) der TSK CIR im GB BMVg ausgebracht wird und die Entwicklerteams mit den im Einsatzraum befindlichen dimensionsspezifischen KI-Systemanteilen verbindet. Über die Bereitstellung einer solchen KI-Infrastruktur zur verzahnten Planung, Datenbereitstellung, Entwicklung, Integration, zum Test und Betrieb solcher Systemanteile ist ein mit der technischen Entwicklung schritthaltender Einsatz der Schlüssel-

³³ Clusterprogramme bilden die Finanzbedarfe in Finanzplanung und Haushalt auf Clusterprogrammebene ab und werden jährlich fortgeschrieben.

technologie KI steuerbar. Diese KI-Infrastruktur muss von zentralen Instanzen bis in die landmobilen, fliegenden und seegehenden Plattformen und zu den Soldatinnen und Soldaten sowie zum „abgesessenen Mitarbeitenden“ reichen. Hierbei wird unterschieden zwischen einer „KI-Infrastruktur für den Einsatz“, über welche entsprechende KI-Systeme und -Modelle zentral bereitgestellt und weiterentwickelt werden können, und einer „KI-Infrastruktur im Einsatz“, welche eine Nutzung der Modelle sowie die Bereitstellung von Daten und Erkenntnissen für eine schnelle und verzugslose Weiterentwicklung der KI-Modelle aus den Erfordernissen des Einsatzes heraus ermöglicht (siehe Abb. 2). Die „KI-Infrastrukturen für den Einsatz und im Einsatz“ umfassen auch die erforderliche Infrastruktur für den Stabs- und Verwaltungsdienst im gesamten GB BMVg.

337. Der zentrale Anteil der KI-Infrastruktur stellt, cloudbasiert, Ressourcen für das Erproben, Entwickeln, Anpassen, Integrieren und/oder Testen von KI-Modellen sowie zur Speicherung von Daten und Modellen zur Weitergabe an die fähigkeitsspezifische Infrastruktur bereit. Dabei muss die Infrastruktur Besonderheiten von eingestufteten Daten und Artefakten mit dadurch eingeschränkter Freigabemöglichkeit sowie entsprechender Ereignisprotokollierung/-auswertung berücksichtigen. Dazu wird das A&S Service Mesh auf der pCloudBw aufgebaut und durch die BWI betrieben werden.

338. Die Nutzung und das (Nach-)Training von KI-Systemen bedingt eine erhebliche Rechenleistung, die wiederum Auswirkungen auf die Rechenzentrumskapazität und die Energieversorgung hat. Dies ist einerseits im Rahmen der Weiterentwicklung des Clusters Infrastructure, Cloud Base and User Equipment (ICU) zu berücksichtigen und bei der Auslegung von Edge-KI Systemen, welche auf Waffenträgern oder in mobiler IT zum Einsatz kommen sollen zu beachten. Das (Nach-)Trainieren erfordert zudem Personal, welches im jeweiligen Projektaufsatz einzuplanen ist.

339. Die dimensionsspezifischen Anteile der KI-Infrastruktur machen Produkte mit erheblichen Datenmengen verfügbar und ermöglichen den Einsatz sowie notwendige Modifikationen der zentral bereitgestellten Modelle. Der Betrieb wird, wo erforderlich, durch fog- bzw. edge-computing-Rechenleistungen skalierbar (dezentrale, cloudfähige Rechenzentren/-leistungen) realisiert.

3.3 Handlungsfeld Verantwortungsvolle Nutzung von Künstlicher Intelligenz

340. Der GB BMVg ist einer verantwortungsvollen Nutzung von KI innerhalb der in dem Abschnitt 2 gesetzten Rahmenbedingungen verpflichtet. Hieraus ergibt sich neben der Schaffung einer klaren Governance-Struktur (siehe Abschnitt 4.1) die Einführung eines RM für KI sowie die Sicherstellung der Umsetzung der wesentlichen Anforderungen an vertrauenswürdige KI gemäß den Prinzipien von NATO und EU. Diese Umsetzung schließt sowohl technische wie auch nicht-technische Maßnahmen ein, um die verantwortungsvolle Nutzung von KI durchgängig sicherzustellen.

3.3.1 Risikomanagement für Künstliche Intelligenz

341. RM KI ist ein Handlungsfeld der KI-Governance. Dieses RM KI muss etabliert werden, um einen sicheren Betrieb zu gewährleisten sowie Schwachstellen, Bedrohungen und Risiken frühzeitig zu identifizieren, zu beheben oder zu mitigieren, bevor es zu Schäden kommt.

342. Die Vorteile des Einsatzes von KI müssen gegen potenzielle Risiken (z. B. Fehlfunktion, Angriffe gegen KI-Infrastrukturen, Datenschutzverletzungen, Sicherheitsrisiken sowie ethische und rechtliche Risiken) abgewogen werden. Hieraus sind entweder Mitigationsmaßnahmen zur Folgenreduzierung bei Eintritt oder – als Ultima Ratio – Verbote abzuleiten.

343. Neben den Risiken des Einsatzes von KI-Systemen müssen beim RM KI auch die Risiken des Nicht-Einsatzes von KI-Systemen berücksichtigt werden. Der Verzicht auf den Einsatz von KI-Systemen kann langfristige und verdeckte Risiken und Kosten und die Unterlegenheit gegenüber feindlichen Kräften mit sich bringen. Angesichts sich verknappender personeller Ressourcen droht der Verlust von Fähigkeiten bzw. die Beeinträchtigung der Aufgabenerfüllung. Oft führen manuelle Prozesse zu erheblichen Effizienzverlusten, da Menschen teilweise eine höhere Fehleranfälligkeit und langsamere Reaktionszeit aufweisen als KI-Systeme. KI kann zur Lösung einiger Herausforderungen signifikant beitragen und dabei aber auch neue Herausforderungen schaffen.

344. Das RM KI betrachtet alle Phasen des Lebenszyklus eines KI-Systems, von der initialen Analyse und Realisierung über die Nutzung bis hin zur Ausmusterung des Systems. Das RM KI unterscheidet dabei grundsätzlich nicht zwischen dem Einsatz von KI-Systemen für die Verwaltung und dem Einsatz für die Streitkräfte.

345. Prinzipiell geben NATO und EU vor, dass das RM KI bestehende Risikomanagementstrukturen ergänzt und in diesen integriert durchzuführen ist. Das in der Realisierungsphase neuer Projekte projektbasierte RM sowie die Informationssicherheit leisten keine gesamtheitliche Betrachtung der Risiken von KI über den gesamten Lebenszyklus von KI-Systemen des GB BMVg und KI-spezifischen IT-Sicherheitsrisiken. Dafür sind Verantwortlichkeiten, Aufgaben und Abläufe gesondert zu regeln.

346. Aus den bestehenden Vorgaben und den strategischen Zielen für die Nutzung von KI (siehe Abschnitt 1), konkret dem Erhöhen der Durchsetzungsfähigkeit, der Steigerung der Effizienz und/oder Effektivität und der Steigerung der Fähigkeit „KI“ im gesamten GB BMVg, ergibt sich folgende **Haltung** für das RM KI im GB BMVg:

„Als GB BMVg wählen wir für KI-Systeme einen interoperablen Risikomanagementansatz, der die positiven Chancen maximiert und die damit einhergehenden negativen Risiken mitigiert. Diese Abwägung wird in den Entscheidungsgremien der KI-Governance ebenengerecht und transparent getroffen, um der Entwicklung von KI „at the speed of relevance“ zu folgen. Oberste Prämisse ist dabei die freiheitlich demokratische Grundordnung als Wertegerüst.“

Generalleutnant Michael Vetter, Ressort CIO und Abteilungsleiter IC im BMVg

Das RM KI des GB BMVg sieht eine Ermittlung, Analyse, Bewertung und Mitigation von Risiken durch konkrete Maßnahmen vor.³⁴ Es muss geprüft werden, ob bestehende Strukturen (z. B. Gremien im Bereich Innovation & Cyber (IC)) und Abläufe oder Gremien der neuen KI-Governance im nachgeordneten Bereich (ngBer) das Aufgabenspektrum des RM KI übernehmen können. Können diese neuen, abgestimmten Aufgaben nicht durch bestehende Rollen und Abläufe aufgefangen werden, müssen neue Rollen und Abläufe im bestehenden Ressourcenrahmen geschaffen werden (z. B. AI Risk Officers oder Lebenszyklus-Abschnitts-Verantwortliche). In der Anlage 5.2 werden weitere Erläuterungen zur Einführung eines RM KI im GB BMVg ausgeführt.

3.3.2 Transparenz, Rückverfolgbarkeit, Dokumentation

3.3.2.1 Transparenz und Nachvollziehbarkeit

347. Transparenz bezeichnet das begrenzte nach außen, vor allem aber nach innen gerichtete Offenlegen relevanter Informationen – von Datenquellen über Modelllogik bis hin zu Einsatzgrenzen – sodass berechnete Interessenstragende den Umgang mit einem KI-System grundsätzlich verstehen können.³⁵ Transparenz ist somit die zentrale Voraussetzung für die Erklärbarkeit von KI-Systemen. Der nach innen und außen gerichtete Transparenzbedarf muss selektiv, nachvollziehbar und adressatengerecht vermittelt werden. Somit wird sichergestellt, dass der bzw. die jeweilige Interessenstragende unter Einhaltung aller sicherheitsrelevanten Vorgaben, alle nötigen Informationen transparent erhält. Nachvollziehbarkeit (Erklärbarkeit) stellt sicher, dass einzelne Entscheidungen oder Ergebnisse technisch wie sachlich begründet und logisch verständlich, erklärbar und dadurch nachvollziehbar sind. Alle verarbeiteten Daten, die Systemlogik und das zugrundeliegende Einsatzmodell werden nachvollziehbar dokumentiert. Die Ergebnisse müssen für Anwendende verständlich und – bei Bedarf – im Detail nachprüfbar sein. Die beteiligten Partner stellen daher umfassende Informationen zur technischen Funktionsweise und zu den Grenzen des Systems bereit. So wird sowohl das Vertrauen der Nutzerinnen und Nutzer gestärkt als auch eine gerichtliche Überprüfung ermöglicht. Für den Bereich

³⁴ EU KI-VO, Art. 9.

³⁵ Siehe DSGVO Artikel 5 Absatz 1 lit. a).

des Verwaltungshandelns ist zu beachten, dass die Nachvollziehbarkeit des Verwaltungshandelns Teil des Rechtsstaatsprinzips ist; dies gilt auch bei Unterstützung durch KI-Systeme. Jedoch sollte jederzeit das „need-to-know“ Prinzip, besonders für sicherheitsrelevante Informationen (z. B. Systemarchitektur) gelten. Somit wird die Transparenz in Bezug auf Sicherheit begrenzt, indem nur berechtigte Interessenstragende, die die entsprechenden Informationen erhalten und nachvollziehen müssen Einblicke erhalten und niemand darüber hinaus diese Informationen erhält.³⁶

3.3.2.2 Technische Sicherung der Nachvollziehbarkeit, Dokumentation und Rückverfolgbarkeit

348. Das Kriterium der Rückverfolgbarkeit richtet den Blick auf die Zeitachse: Alle Daten-, Modell- und Prozessschritte sind lückenlos rekonstruierbar, sodass sich jede Systemversion und jeder Entscheidungsweg im Detail zurückverfolgen lassen. Die hierfür relevanten Komponenten sind:

- **Ereignisprotokollierung:** Automatisierte Audit-Trails erfassen Datenzugriffe (Input, Prompt), Systemänderungen und KI-Entscheidungen (Output) in einem unveränderbaren Speicher.
- **Versionstransparenz:** Alle Modell-, Daten- und Konfigurationsstände erhalten eindeutige Referenzen, sodass frühere Zustände jederzeit rekonstruierbar bleiben.
- **Kontinuierliche Entscheidungspfad aufzeichnung:** Bei sicherheitskritischen Anwendungen werden alle Entscheidungspfade lückenlos aufgezeichnet. Dennoch ist es wichtig zu beachten, dass Transparenz nicht grenzenlos sein darf. Auch hier gilt das „need-to-know“ Prinzip. So können gewisse Informationen nicht offengelegt werden und müssen vertraulich gegenüber Externen dokumentiert werden. Somit bleiben sicherheitskritische Details eingeschränkt. Transparenz soll dieses Blackboxing aufbrechen, soweit es kein Sicherheitsproblem erzeugt und sicherheitsrelevante Daten offenlegt.³⁷

349. Dokumentation ist die formalisierte, dauerhafte Aufzeichnung aller genannten Informationen; sie schafft die strukturierte Grundlage, auf der Transparenz ermöglicht, Nachvollziehbarkeit demonstriert und Rückverfolgbarkeit technisch umgesetzt werden kann (siehe Abb. 3). Sämtliche Entwicklungs- und Bewertungsprozesse werden systematisch und – soweit möglich – automatisiert dokumentiert, um Transparenz sowie Rückverfolgbarkeit zu gewährleisten. Jedes KI-System verfügt über eine Begleitdokumentation, die Zweck, Leistungsparameter, Einsatzgrenzen sowie Risiko- und Bias-Analysen bündelt. Änderungen in Entwicklung, Training, Inbetriebnahme und Betrieb aktualisieren die Dokumentation automatisch und schaffen einen belastbaren Nachweis für Audits.

³⁶ Transparenz von KI-Systemen BSI Whitepaper.

³⁷ Transparenz von KI-Systemen BSI Whitepaper.

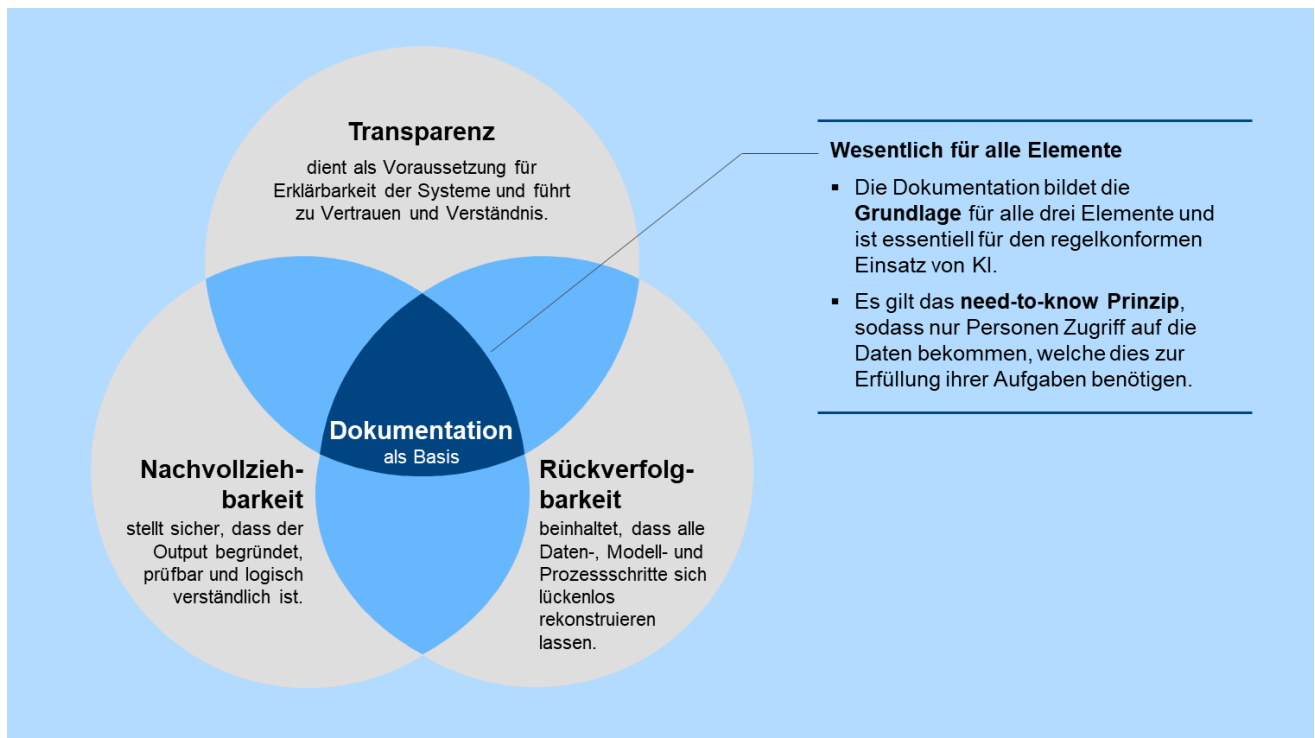


Abb. 2: Abgrenzung der Themen Transparenz, Nachvollziehbarkeit & Rückverfolgbarkeit

3.3.2.3 Prozessuale Anforderungen

350. Für den Erhalt des Kontroll- und Vertrauensgerüsts bedarf es einiger prozessualer Festlegungen:

- **Rollenbasierte Verantwortungszuordnung:** Für jede Entscheidung ist eindeutig eine verantwortliche Person oder Systemkomponente benannt.
- **Kontextabhängige Erklärbarkeit:** Die Tiefe der Erklärung richtet sich nach Einsatzrisiko und menschlichem Entscheidungsspielraum.
- **Regelmäßige Prüfprotokolle:** Zyklische Reviews sichern Funktionsfähigkeit, Leistungsstabilität und Integrität aller Nachweise über den gesamten Lebenszyklus hinweg – einschließlich Archivierung oder Löschung.

3.3.3 Trainieren, Verifizieren, Validieren und Testen

3.3.3.1 Trainieren der Künstlichen Intelligenz

351. Beim maschinellen Lernen haben die während des Trainings verwendeten Daten einen erheblichen Einfluss auf das spätere Verhalten des Modells. Aus diesem Grund müssen Trainings-, Validierungs- und Testdaten mit deren Metadaten sowie Modelleinstellungen (Hyperparameter) revisionsgerecht dokumentiert, zentral gepflegt und gemäß eines KI-Lebenszyklus u. a. für ein potenziell erforderliches Nachtraining verwaltet werden. Dies ermöglicht eine Wiederholung des Trainings unter Erzeugung vergleichbarer Ergebnisse.

352. Ein Nachtraining eines KI-Systems muss erfolgen, wenn sich die zugrundeliegenden Daten, Einsatzvorgaben, Fähigkeitsanforderungen oder regulatorischen Rahmenbedingungen wesentlich verändert haben oder wenn im laufenden Betrieb systematische Fehlverhalten, Leistungsabfall oder unerwünschte Verzerrungen festgestellt werden.

353. Ein Nachtraining ist auch dann notwendig, wenn neue, qualitätsgesicherte Daten vorliegen, die eine signifikante Verbesserung der Modellleistung erwarten lassen. Datensets müssen grundsätzlich qualitätsgesichert, versioniert und dokumentiert werden.

354. Qualitätssicherung umfasst dabei sowohl die Vorverarbeitung der Daten zur besseren Verwendbarkeit als auch die Bestimmung und Dokumentation von Eigenschaften und statistischen Kennzahlen der Datensets. Es soll rückverfolgbar sein, welche Daten in welcher Version in das Training oder die Weiterentwicklung eines KI-Modells geflossen sind. Diese sollen dauerhaft verfügbar sein. Die während des Trainings verwendeten Daten sollen aus einer möglichst großen Zahl unabhängiger, vertrauenswürdiger Quellen mit glaubhaften und kontrollierten Inhalten (Qualität) zusammengestellt werden.

3.3.3.2 Verifizieren und Validieren

355. Um sicherzustellen, dass die Ausgaben eines KI-Modells Vorgaben und Spezifikationen erfüllen, kommen Verifikations- und Validierungstechniken (V&V-Techniken) zum Einsatz. Bei der Verifikation wird überprüft, ob ein KI-Modell gemäß den angegebenen Anforderungen erstellt und entwickelt wurde, wohingegen bei der Validierung überprüft wird, ob das Produkt die Anforderungen und Erwartungen erfüllt. V&V-Methoden ermöglichen darüber hinaus eine frühzeitige Erkennung von Bugs und Strategien zum Umgang mit verzerrten Daten (Data Bias).

356. Die Validierung von KI-Systemen stellt sicher, dass diese im vorgesehenen Anwendungskontext verlässlich, robust und regelkonform arbeiten. Vor dem Einsatz müssen KI-Systeme daher einer strukturierten Validierung unterzogen werden, die unter vorab definierten Bedingungen prüft, ob die erwartete Funktionalität und Leistung gemäß den zuvor festgelegten Systemanforderungen erreicht wird. Die Validierung umfasst insbesondere die Überprüfung auf Genauigkeit und Robustheit gegenüber Eingabeänderungen sowie das Verhalten des Systems in Fehlersituationen. Sie erfolgt mit einem separaten, qualitativ hochwertigen und repräsentativen Validierungsdatensatz, der nicht für das Training verwendet wurde. Dabei sind auch im Rahmen des RM (siehe Abschnitt 3.3.1) potenzielle Verzerrungen sowie Risiken für Grundrechte, Gesundheit, Sicherheit und Umwelt zu bewerten.

357. Der Validierungsprozess ist dokumentationspflichtig und muss reproduzierbar sowie nachvollziehbar gestaltet sein. Alle durchgeführten Validierungsschritte, inklusive der genutzten Datensätze, Metriken, Ergebnisse und Bewertungen, sind zentral zu erfassen und versioniert zu speichern. Um möglichst objektive Bewertungen zu gewährleisten, muss zusätzlich sichergestellt werden, dass die Validierung durch qualifiziertes Fachpersonal erfolgt, welches nicht unmittelbar an der Entwicklung des KI-Systems beteiligt war (Vier-Augen-Prinzip). Änderungen am System, etwa durch das oben beschriebene Nachtraining, erfordern eine erneute Validierung gemäß denselben Anforderungen.

3.3.3.3 Testen

358. Die Leistungsfähigkeit des KI-Systems wird mit repräsentativen Testdaten umfassend geprüft und ist Voraussetzung für die Erteilung einer Genehmigung zur Nutzung (GeNu – siehe Abschnitt 3.3.4.2).

359. Es müssen folgende Testdatensätze spezifisch für jedes durch die Bw entwickelte oder weiterentwickelte KI-System erzeugt, verwaltet und verfügbar gemacht werden:

1. Zero-Failure-Tests stellen, im Fall eines KI-Systems mit hohem Risiko, eine notwendige Minimalanforderung an das zu testende KI-Modell dar.
2. Qualitätstests sind zur Messung der Güte des KI-Modells zu verwenden. Beispielhafte Metriken sind Accuracy oder F1-Score des Modells.

360. Beide Typen von Testdatensätzen sollen in festgelegten Abständen aktualisiert und dabei einem konsistenten Schema folgend versioniert werden. Trainings-, Validierungs- und Testdaten sind dauerhaft getrennt zu halten. Testdaten dürfen weder für das Training noch die Validierung verwendet worden sein.

3.3.3.4 Vorurteilsminderung

361. Im Rahmen von Trainings-, Validierungs- und Testabläufen eines KI-Systems können sowohl die dort verwendeten Datensätze als auch die, die im Einsatz des Systems verarbeitet werden können, unabsichtlich zu Verzerrungen führen. Etwaige Verzerrungen können schlimmstenfalls fortgeschrieben werden und sich dadurch weiter verschärfen. Hieraus können falsche oder verzerrte Ergebnisse resultieren, es kann sich aber auch eine Diskriminierung ergeben, die bestimmte Personengruppen benachteiligt; hiervon abzugrenzen sind Anforderungen, die sich aus dem militärischen Aufklärungserfordernis ergeben können. Für den verantwortungsvollen Einsatz von KI ergibt sich damit das unbedingte Erfordernis, dass Fairness und unintendierte Diskriminierung als integrale Bestandteile im Umgang mit KI verstanden und somit immer mitberücksichtigt werden.

362. Diese Maßnahmen ergänzen die bestehenden Grundsätze der Bw um KI-spezifische Anforderungen. Sie sind als verbindlicher Mindeststandard in Leistungsbeschreibungen, Verträgen und Abnahmeprozessen zu verankern und unterliegen unabhängigen Prüfungen.

3.3.4 Konformitätsprüfung, Genehmigung zur Nutzung und Inbetriebnahme

3.3.4.1 Erforderliche Funktionalitäten zur Inbetriebnahme von KI-Systemen

363. Für die Inbetriebnahme von KI-Systemen sind Funktionalitäten für das Deployment auf Plattformen der Bw, für das Performance Monitoring und für (in-mission) Updates erforderlich. Dies beinhaltet das Einbringen neuer Fähigkeiten durch „in-mission Model Deployment-Fähigkeiten“ auf Plattformen, das Nachtrainieren von Systemen anhand neuer Daten sowie darauf basierende in-mission Updates nach erfolgreicher (Re-)Validierung der neuen Systeme. Zum Deployment zählt weiterhin ein automatisiertes in-mission Monitoring der System-Ergebnisse sowie ein zentrales Management der eingesetzten Systeme über alle Plattformen.

364. Bei der Entwicklung von KI-Systemen bevorzugt der GB BMVg Organisationen und Unternehmen, die durch nationale Stellen, EU und/oder NATO zertifiziert sind und sich zur Einhaltung rechtlicher sowie ethischer Standards (siehe Abschnitte 2.2, 2.3 und 2.4, hier z. B. NATO Prinzipien für verantwortungsvollen Umgang) bekennen.

3.3.4.2 Konformitätsprüfung, Genehmigung zur Nutzung

365. Die Festlegung der Anforderungen und die erforderlichen Funktionalitäten zur Inbetriebnahme von KI-Systemen sowie das Testen mittels Testdaten (siehe Abschnitt 3.3.3.3) sind die Vorausset-

zungen dafür, dass eine Konformitätsprüfung durchgeführt werden kann. Eine Konformitätsprüfung im Kontext von KI bezieht sich auf die Überprüfung, ob ein KI-System die relevanten Gesetze, Vorschriften, die „Principles of Responsible Use der NATO“, ethischen Normen sowie demokratischen Grundwerte einhält, die für die Entwicklung, Implementierung und Betrieb von KI-Systemen gelten. Die Konformitätsprüfung ist Teil der integrierten Nachweisführung vor GeNu nach der Projektbezogenen Bedarfsdeckung und Nutzung (PBN) und wichtig, um sicherzustellen, dass KI-Systeme sicher, verantwortungsvoll und gesetzeskonform eingesetzt werden. Eine Konformitätsprüfung muss vor dem ersten Inverkehrbringen eines KI-Systems, sowie nach wesentlichen Änderungen des KI-Systems erfolgen. Der GB BMVg orientiert sich bezüglich der Ausgestaltung der Konformitätsprüfung und GeNu von KI-Systemen an den Vorgaben der NATO bzw. der EU KI-VO. Die erfolgreiche bestandene Konformitätsprüfung und das Vorliegen der GeNu sind grundlegende Voraussetzungen für die Inbetriebnahme von KI-Systemen im GB BMVg. Die rechtlich vorgegebene Waffenprüfung bei entsprechenden Systemen findet hierzu eigenständig statt.

3.3.5 Menschliche Kontrolle

366. Es gibt noch kein einheitliches internationales Verständnis, wie menschliche Kontrolle (siehe Abschnitte 2.1 und 2.3) konkret zu operationalisieren ist. DEU setzt sich gemeinsam mit FRA dafür ein, dass sich alle CCW-Vertragsstaaten³⁸ in einem sogenannten „normativen und operativen Rahmenwerk“ („normative and operational framework“) verpflichten, gewisse Parameter zur Sicherstellung menschlicher Kontrolle zu implementieren.

367. Aus DEU und französischer Sicht sind Kontext bezogenes, sinnvolles menschliches Ermessen und menschliche Kontrolle („context-appropriate human judgement and control“)³⁹ im Rahmen einer verantwortlichen Befehlskette gewährleistet, wenn Menschen mit Blick auf WaSys mit autonomen KI-Funktionen u. a.

- angemessen ausgebildet werden,
- deren Funktionsweise, Effekte und Interaktion mit Umgebungseinflüssen verstehen,
- deren Zuverlässigkeit evaluieren und beobachten können,
- deren Einsatzregeln festlegen und validieren können,
- einen präzisen Rahmen für deren Missionen festlegen, inklusive Missionsziel, Zieltypen, Missionsdauer und Einsatzgebiet,
- deren Zuverlässigkeit, Funktions- und Betriebssicherheit sowie Vorhersagbarkeit durch Qualifikation (Testen, Zertifizierung) und ungewolltem Zugriff sicherstellen,

³⁸ Certain Conventional Weapons.

³⁹ Siehe Paris Declaration on Maintaining Human Control in AI-enabled Weapon Systems. Élysée. (<https://www.elysee.fr/emmanuel-macron/2025/02/11/paris-declaration-on-maintaining-human-control-in-ai-enabled-weapon-systems> (abgerufen am 17. August 2025))

- rechtliche Waffenprüfverfahren gemäß Artikel 36, I. Zusatzprotokoll zu den Genfer Abkommen durchführen,
- die Verantwortung für die Entscheidung über deren Einsatz und Einsatzparameter sowie im Fall der Fälle auch für Verletzungen des Humanitären Völkerrechts mit solchen WaSys tragen und
- solche WaSys nicht außerhalb menschlicher Befehlsketten und nur im Rahmen verantwortungsvoller Befehlsketten einsetzen.

368. Der GB BMVg muss sicherstellen, dass kontextbezogenes, menschliches Urteilsvermögen und menschliche Kontrolle bei WaSys mit autonomen Funktionen im Rahmen einer verantwortlichen Befehlskette gewährleistet und sichergestellt sind.

3.3.6 Zuverlässigkeit, Robustheit, Sicherheit und Datenschutz

3.3.6.1 Fortlaufende Risikoüberwachung und Performancebewertung

369. Um die Wirksamkeit der KI-Systeme im realen Einsatz sicherzustellen, werden regelmäßige Überprüfungen der Systemleistung durchgeführt. Die laufende Risikoüberwachung im Rahmen des RM KI ermöglicht es, auf mögliche Schwachstellen und unvorhergesehene sowie unintendierte Auswirkungen frühzeitig zu reagieren. Hierbei spielen verpflichtende Kennzahlen wie Genauigkeit, Zuverlässigkeit und Fehlerraten eine zentrale Rolle, um sicherzustellen, dass jedes System den festgelegten Qualitätsstandards entspricht.

3.3.6.2 Feedback und kontinuierliche Anpassung

370. Basierend auf Feedback und den in der Risikoüberwachung identifizierten Anpassungserfordernissen werden die KI-Systeme kontinuierlich verbessert. Diese iterative Anpassung stellt sicher, dass die Technologie den dynamischen Anforderungen im bundeswehrspezifischen Kontext während der gesamten Nutzungsdauer gerecht wird. Wichtig ist, dass vor allem bei WaSys die iterative Anpassung in einer relativ hohen Frequenz nach neuesten technischen Standards und Gefechtsfeldanforderungen erfolgt.

3.3.6.3 Revision und Auditierung

371. Alle Aktivitäten und Anpassungen in der Nutzung der Systeme werden genau und revisions-sicher archiviert, um die Rückverfolgbarkeit und die Nachvollziehbarkeit der Ergebnisse zu gewährleisten. Regelmäßige Audits⁴⁰ stellen sicher, dass die Systeme auch im laufenden Betrieb den ethischen, rechtlichen und technischen Standards entsprechen.

⁴⁰ ISO/IEC 5259-3:2024 „Künstliche Intelligenz — Datenqualität für Analytik und maschinelles Lernen (ML) — Teil 3: Anforderungen und Leitlinien für das Datenqualitätsmanagement“.

3.3.6.4 Datenschutz

372. Im Kontext von KI ist der Schutz von personenbezogenen Daten von besonderer Bedeutung. KI-Systeme müssen während ihres gesamten Lebenszyklus die datenschutzrechtlichen Anforderungen angemessen einhalten. Entsprechend dem Grundsatz der Datenminimierung dürfen nur solche personenbezogenen Daten verarbeitet werden, die für den jeweiligen legitimen Zweck zwingend erforderlich sind. Daten, die unter „besondere Kategorien personenbezogener Daten“⁴¹ fallen (z. B. ethnische Herkunft, religiöse Überzeugungen), unterfallen speziellen Verarbeitungsanforderungen.

373. Grundsätzlich sollte bei der Entwicklung von Systemen das Lernen mit anonymisierten oder pseudonymisierten Daten bevorzugt werden, da die Vorgaben des Datenschutzes nur auf personenbezogene (also nicht anonyme) Daten Anwendung finden. Daher sollten Anonymisierungs-, Pseudonymisierungs- und Synthetisierungsverfahren für Trainings- und Testdaten angewandt werden. Anonymisiert können die Daten jedoch von KI-Systemen weiterverwendet werden.⁴² Originaldaten werden so weit wie nötig verschleiert und dann gelöscht; die Auswertung erfolgt nur auf Basis der verschleierten Daten.

374. Vor der Einführung eines KI-Systems muss eine Datenschutz-Folgenabschätzung durchgeführt werden, um mögliche Risiken für die Privatsphäre der Betroffenen zu identifizieren und geeignete Maßnahmen zur Risikominimierung festzulegen. Technisch-organisatorische Maßnahmen sind anzuwenden, um die Einhaltung des Datenschutzes bei KI-Systemen sicherzustellen.

375. Grundsätzlich müssen alle gesetzlichen Vorgaben zum Datenschutz berücksichtigt werden. Um eine hohe Datenverfügbarkeit und Datenqualität zu gewährleisten, müssen jedoch Ausnahmen im Rahmen des Datenschutzes für militärische Zwecke geprüft und konsequent genutzt werden.

3.3.6.5 Daten- und Informationssicherheit

376. Die Vorgaben zur Daten- und Informationssicherheit sind der A-960/1 zu entnehmen. Sicherheitsmaßnahmen sollen darauf abzielen, das Risiko eines Missbrauchs oder einer Schädigung zu verringern und Sicherheitsverletzungen oder Systemausfälle zu verhindern. Daten sollen in Systemen gespeichert werden, die gegen Angriffe von außen und innen geschützt sind. Der Zugriff der KI auf die Daten soll strikt auf die vorgesehenen Verwendungszwecke beschränkt sein. Um sicherzustellen, dass nur befugte Personen Zugriff haben, müssen sowohl bei der Speicherung als auch bei der Übertragung von Daten Sicherheitsmechanismen (z. B. Firewalls, Multi-Faktor-Authentifizierung) eingerichtet werden.

⁴¹ DSGVO Artikel 9.

⁴² Nach Auffassung der bzw. des Bundesbeauftragten für Datenschutz und Informationssicherheit (BfDI) bedarf die Anonymisierung als Verarbeitungsschritt allerdings einer Rechtsgrundlage. Je nach Kontext kann diese über die Zweckänderungsvorschriften des BDSG begründet werden.

4 KI-Governance und Umsetzungsschritte

4.1 KI-Governance im Geschäftsbereich des BMVg

401. Für die effektive und zielgerichtete Regulierung, Steuerung und Überwachung der Entwicklung und des Einsatzes von KI im GB BMVg ist eine KI-Governance⁴³ essenziell, um den verantwortungsvollen, sicheren und regelkonformen Einsatz von KI zu gewährleisten. Dies gilt insbesondere im Hinblick auf ethische Standards, die Konsistenz des Einsatzes von KI-Systemen innerhalb des GB BMVg sowie die Einhaltung regulatorischer Anforderungen und die Erfüllung der im Rahmen der NATO eingegangenen Verpflichtungen.

402. Die Notwendigkeit einer KI-Governance ergibt sich wesentlich aus den Vorgaben der EU KI-VO und der NATO KI-Strategie. Die KI-Governance, das Chief Artificial Intelligence Officer Council (CAIO-Council) und die Integrierten Aufgabenteams (IAT) sind dabei keine ministeriellen Arbeitsgruppen im Sinne der ergänzenden Geschäftsordnung des BMVg. Entscheidungsbefugnisse und Verantwortlichkeiten erfolgen weiterhin unter Berücksichtigung der Zuständigkeiten für die (Geschäfts-)Prozesse. Die KI-Governance bringt Entscheidungstragende einschließlich Prozessverantwortliche zielgerichtet zur Abstimmung und Beratung zusammen, um Maßnahmen zur Nutzung von KI im GB BMVg im Sinne ihres Auftrages voranzutreiben. So kann eine zentral koordinierende, handlungsfähige, reaktionsschnelle und anschlussfähige Steuerung von KI im GB BMVg erreicht werden, ohne in bestehende Verantwortlichkeiten einzugreifen. Die KI-Governance ersetzt nicht bestehende Geschäftsverteilungen, Zuständigkeiten, Prozesse und Verfahren.

403. Die KI ist ein Themenfeld in der Digitalisierung des GB BMVg. KI-spezifische Aufgaben werden daher, wo möglich, durch die derzeitige Struktur zur Steuerung der Digitalisierung im GB BMVg mit abgedeckt. Der Ansatz der zentralen Steuerung bei dezentraler Umsetzung wird auch mit diesem KD verfolgt. Die KI-Governance nutzt für den breiten Austausch zudem das bestehende Netzwerk Digitalisierung mit den Koordinatorinnen und Koordinatoren Digitalisierung. Das Veränderungsmanagement KI erfolgt zusammen mit dem etablierten Veränderungsmanagement in der Digitalen Transformation, um Doppelstrukturen zu vermeiden. Der durch die Koordinatorin bzw. den Koordinatoren bestimmte Point of Contact (POC) Digitalisierung oder der bedarfsweise ergänzend bestimmte POC KI fungiert für den eigenen Bereich als Ansprechperson im Themenfeld KI und stellt die Kommunikation und Information in den jeweiligen Bereichen sicher.

404. Die (Haupt-)Prozessverantwortlichen sind in der Maßnahmenumsetzung wesentlich und dementsprechend in die IAT-Arbeit einzubinden. Die Verantwortung für die eigenen Verfahren, für die

⁴³ Governance bezeichnet allgemein ein Steuerungs- und Regelungssystem einer Organisation; es gibt hierzu keine deutsche Entsprechung. Die KI-Governance hat dementsprechend keine eigenen organisatorischen Aufgaben sondern dient der Koordination und Steuerung.

dort verarbeiteten Informationen und das dazu benötigte Wissen verbleibt bei der jeweiligen Fachlichkeit.

405. Die KI-Governance ordnet sich als eine Governance in den Gesamtverantwortungsbereich der bzw. des Ressort-CIO ein. Eine künftige Zusammenführung der KI-Governance mit der Data Governance ist im Rahmen einer Überarbeitung der Datenstrategie grundsätzlich anzustreben. Entwicklungen in der NATO sind zu berücksichtigen.

406. Die bzw. der CAIO des GB BMVg steuert das Thema KI für den gesamten GB BMVg auf strategischer Ebene. Die bzw. der CAIO des GB BMVg übernimmt die Leitung des ablauforganisatorisch eingerichteten CAIO-Councils. Ihre bzw. seine Aufgabe im Rahmen der KI-Governance ist es, auf strategischer Ebene für eine zielgerichtete und bedarfsgerechte Einführung, Weiterentwicklung und Nutzung von KI im gesamten GB zu sorgen. Dazu gehört insbesondere das

- Begleiten und Koordinieren der Umsetzung der KI-Themenstellungen entlang der Roadmap auf strategischer Ebene,
- kontinuierliches Analysieren und Prüfen von Einsatzpotenzialen von KI auf strategischer Ebene,
- Beraten der Vorhaben und Unterstützen bei Entscheidungsbedarfen auf strategischer Ebene,
- Einnehmen einer Vorbildrolle im eigenen Verantwortungsbereich und darüber hinaus, um Orientierung und Handlungssicherheit zu schaffen,
- Pflegen und Ausbauen von Kontakten in die Fachwelt in Forschung, Wirtschaft und Gesellschaft,
- Vertreten der nationalen Position in internationalen Gremien der Cyberpolitik sowie in der NATO und der EU sowie
- Umsetzen und kontinuierliches Weiterentwickeln der KI-Governance.

407. Die bzw. der CAIO der Bw koordiniert operativ das Thema KI in der Bw. Die oder der CAIO der Bw ist Mitglied im CAIO-Council. Seine bzw. ihre rollenspezifische Aufgabe im Rahmen der KI-Governance ist es, für eine zielgerichtete und bedarfsgerechte Einführung, Weiterentwicklung und Nutzung des Themenfelds KI in der Bw zu sorgen. Dazu gehört insbesondere das

- Begleiten und Koordinieren der Realisierung von KI-Aktivitäten⁴⁴ der Bereiche der Bw^{45, 46},
- Unterstützen bei der kontinuierlichen Analyse und dem Prüfen von Einsatzpotenzialen von KI durch die Bereiche und die Prozessverantwortlichen,

⁴⁴ Unter KI-Aktivitäten sind sowohl die Anteile der Bereiche an den Aufgaben der KI-Governance sowie grundsätzlich die Tätigkeiten und Projekte mit KI(-Bezug) in den Bereichen zu verstehen.

⁴⁵ Als Bereiche werden hier die Militärischen Organisationsbereiche (MilOrgBer) und zivilen OrgBer sowie die dem BMVg unmittelbar nachgeordneten DSt betrachtet.

⁴⁶ Dies erfolgt in Abstimmung mit den Koordinatorinnen und Koordinatoren Digitalisierung und den Prozessverantwortlichen der Bw.

- Beraten der Vorhaben und Unterstützen bei Entscheidungsbedarfen in den Bereichen,
- Einnehmen einer Vorbildrolle im eigenen Verantwortungsbereich, um Orientierung und Handlungssicherheit zu schaffen und
- Pflegen und Ausbauen von Kontakten in die Fachwelt in Forschung, Wirtschaft und Gesellschaft.

408. Ziel des CAIO-Councils ist es, die verantwortungsvolle Nutzung von KI-Systemen im GB BMVg zu ermöglichen und KI-Aktivitäten im GB BMVg zu koordinieren. Das CAIO-Council legt Prinzipien und Leitlinien zur Nutzung von KI fest, erarbeitet KI-spezifische Strategien und Vorschriften, erstellt kontinuierlich ein Lagebild KI und identifiziert und adressiert Herausforderungen im Kontext KI. Darüber hinaus identifiziert es KI-spezifische Themen für das Leitungsboard Digitalisierung und das Steuerungspanel Digitalisierung.

5 Anlagen

5.1	Bereichsausnahme von der EU KI-Verordnung	38
5.2	Erläuterung zur Einführung eines Risikomanagements Künstlicher Intelligenz im Geschäftsbereich des BMVg	45
5.3	Abkürzungsverzeichnis	51

5.1 Bereichsausnahme von der EU KI-Verordnung

5.1.1 Stellungnahme BMVg vom 30. Januar 2025 zur Reichweite der Bereichsausnahme

I. Problemstellung

Im August 2024 ist die Verordnung (EU) 2024/1689, die sogenannte KI-VO in Kraft getreten. Sie nimmt mit Bezug auf KI eine Risikoeinteilung vor, unterteilt diese in verbotene Systeme (Art. 5 KI-VO), Hochrisiko-Systeme (Art. 6 KI-VO + Annex III), sowie Systeme mit normalem und minimalem Risiko. Je nach Risikostufe sieht die VO einen breiten Pflichtenkatalog vor; diese reichen von einem kompletten Verbot über Offenlegungs- und Dokumentationserfordernisse, Durchführung einer Grundrechte-Folgenabschätzung bis hin zu vergleichsweise harmlosen Transparenzpflichten. Diese Pflichten treffen primär Hersteller/Anbieter/Einführer, aber auch bloße Betreiber der KI⁴⁷. Für den Auftrag der Streitkräfte sind insbesondere die Offenlegungspflichten unter Geheimhaltungsaspekten problematisch. Jedoch stellen auch die Dokumentations- und Transparenzpflichten nicht zu unterschätzende bürokratische Hürden auf, welche den erforderlichen schnellen Aufwuchs an Fähigkeiten behindern. Es gilt daher zu prüfen, inwieweit die in Art. 2 (3) KI-VO normierte Bereichsausnahme die Streitkräfte vom Pflichtenprogramm der KI-VO ausnimmt.

II. Gesetzestext

Art. 2 Abs. 3 UAbs. 1-3 KI-VO:

[1] Diese Verordnung gilt nur in den unter das Unionsrecht fallenden Bereichen und berührt keinesfalls die Zuständigkeiten der Mitgliedstaaten in Bezug auf die nationale Sicherheit, unabhängig von der Art der Einrichtung, die von den Mitgliedstaaten mit der Wahrnehmung von Aufgaben im Zusammenhang mit diesen Zuständigkeiten betraut wurde.

[2] Diese Verordnung gilt nicht für KI-Systeme, wenn und soweit sie ausschließlich für militärische Zwecke, Verteidigungszwecke oder Zwecke der nationalen Sicherheit in Verkehr gebracht, in Betrieb genommen oder, mit oder ohne Änderungen, verwendet werden, unabhängig von der Art der Einrichtung, die diese Tätigkeiten ausübt.

[3] Diese Verordnung gilt nicht für KI-Systeme, die nicht in der Union in Verkehr gebracht oder in Betrieb genommen werden, wenn die Ausgaben in der Union ausschließlich für militärische Zwecke, Verteidigungszwecke oder Zwecke der nationalen Sicherheit verwendet werden, unabhängig von der Art der Einrichtung, die diese Tätigkeiten ausübt.

Von Bedeutung in diesem Kontext auch Erwägungsgrund (EG) 24 (siehe am Ende des Dokuments).

⁴⁷ Je nach Konstellation ist durchaus damit zu rechnen, dass die Anbieter-Eigenschaft im Sinne des Artikels 3 Nr. 3 KI-VO vorliegt.

III. Kurzfassung

Artikel 2 Abs. 3 UAbs. 2 KI-VO sieht vor, dass KI-Systeme, wenn und soweit sie **ausschließlich** für *militärische Zwecke*, *Verteidigungszwecke* oder *Zwecke der nationalen Sicherheit* in Verkehr gebracht, in Betrieb genommen oder verwendet werden, vom Geltungsbereich der Verordnung ausgenommen sind (sog. Bereichsausnahme KI-VO).

Bislang existiert keine Rechtsprechung zur KI-VO und folglich auch keine Rechtsprechung zur Reichweite dieser Bereichsausnahme. Allerdings werden in der Rechtsprechung des EuGH zu anderen europäischen Vorschriften die genannten Tatbestände als Ausnahmetatbestände eng ausgelegt.

Eine im GB BMVg eingesetzte KI ist daher nicht per se von der Bereichsausnahme erfasst, nur weil diese im GB BMVg zur Anwendung kommt. Es muss vielmehr einer der in Art. 2 Abs. 3 UAbs. 2 KI-VO aufgeführten Tatbestände (ausschließlich für militärische Zwecke, Verteidigungszwecke oder Zwecke der nationalen Sicherheit, s. o.) erfüllt sein.

Gleichzeitig muss dem Sinn und Zweck der Bereichsausnahme, zu verhindern, dass die Verteidigungsfähigkeit und Sicherheit der Mitgliedstaaten der Union durch etwaige konterkarierende Vorschriften der KI-VO unterminiert wird, Rechnung getragen werden.

Viele Anwendungsfälle von KI, mit denen kurz- und mittelfristig im GB BMVg zu rechnen ist, dürften zwar einen militärischen oder Verteidigungszweck verfolgen und somit unter die Bereichsausnahme des Art. 2 Abs. 3 UAbs. 2 KI-VO fallen. Ebenso dürften Pflichten, welche die Hersteller von Dual-use-Anwendungen treffen, die Streitkräfte dann nicht mehr treffen, sobald sie diese für ihren genuin militärischen Auftrag einsetzen. Dennoch kommt es in Grenzfällen auf eine fundierte Begründung im Einzelfall an. Diese muss aus der Praxis heraus durch die Betreiber (Art. 3 Nr. 4 KI-VO bzw. EG 13) selbst erfolgen. Vorliegende Kommentierung soll hierzu als Leitfaden dienen.

IV. Im Einzelnen

Zu den einzelnen Tatbeständen des Art. 2 Abs. 3 UAbs. 2 KI-VO existiert noch keine Rechtsprechung. Es stellt sich umso mehr die Frage nach der Reichweite dieser Ausnahme. Die Bereichsausnahme ist nicht bereits dadurch gegeben, dass ein KI-System im GB BMVg eingesetzt wird, sondern dass einer der in Artikel 2 Abs. 3 UAbs. 2 KI-VO aufgeführten Tatbestände erfüllt ist:

1. Die in Art. 2 Abs. 3 UAbs. 2 KI-VO genannten Ausnahmetatbestände knüpfen an den **ausschließlichen Verwendungszweck an**, nicht an die jeweilige anwendende Einrichtung. Unabhängig von der Art der Einrichtung, die die Tätigkeit ausübt, kommt es im Einzelfall darauf an, ob ein KI-System/ KI-Modell, das konkret im GB BMVg genutzt wird, den Zwecken Militär, Verteidigung, nationale Sicherheit und/oder z. B. Forschungszwecken (siehe hierzu Art. 2 Abs. 6 und 8, bzw. EG 25) dient.

2. Welcher Zweck nun konkret einen **militärischen**, einen **Verteidigungszweck** oder einen **Zweck der nationalen Sicherheit** abbildet, ist im Lichte der Rechtsprechung des EuGH auszulegen. Dieser legt insbesondere den Begriff der nationalen Sicherheit eng aus:

- a. Danach erfasst **die nationale Sicherheit** die Verhütung und Repression von Tätigkeiten, die geeignet sind, die tragenden Strukturen des Landes im Bereich der Verfassung, Politik oder Wirtschaft oder im sozialen Bereich in schwerwiegender Weise zu destabilisieren und insbesondere die Gesellschaft, die Bevölkerung oder den Staat als solchen unmittelbar zu bedrohen (EuGH Verb. Rs. C-511/18, 512/18 und 520/18 (La Quadrature du Net), ECLI:EU:C:2020:791, Rn. 135).

Umfasst vom Begriff der nationalen Sicherheit sind Tätigkeiten der Inlands- und Auslandsnachrichtendienste sowie Tätigkeiten zur Gewährleistung der Sicherheit des Staates. Es geht im Kern um die Zuständigkeit der Mitgliedstaaten für die Abwehr äußerer Angriffe und innerer Attacken auf die staatliche Ordnung. Letzteres umfasst insbesondere Verhütung und Repression von terroristischen Aktivitäten. Nicht vom Begriff der nationalen Sicherheit erfasst ist der Bereich der Strafverfolgung, einschließlich der Bekämpfung auch schwerer Kriminalität, der eher unter den Begriff der „öffentlichen Sicherheit“ zu fassen ist. Ebenso wenig erfasst sind die Bereiche Migration, Asylwesen und Grenzkontrollmanagement, wie die Aufzählungen der Hochrisiko-KI-Systeme in Annex III Nr. 6 und Nr. 7 belegen (vgl. Martini/Wendehorst, 1. Aufl. 2024, KI-VO Art. 2 Rn. 63).

Ob eine KI i. S. d. Verordnung der Verhütung und Repression der genannten unmittelbaren Bedrohungen dient, ist individuell zu prüfen. Ist dies nicht der Fall, ist dieser Tatbestand der Bereichsausnahme aus Art. 2 Abs. 3 UAbs. 2 KI-VO h.E. nach nicht einschlägig.

- b. Im GB BMVg dürfte häufiger die Fallgruppe „Einsatz einer KI **zu militärischen Zwecken**“ einschlägig sein.

Militärische Zwecke setzen danach voraus, dass der betroffene Gegenstand für den militärischen Einsatz verwendet werden soll (Burgi/Dreher/Opitz/von Wietersheim, 4. Aufl. 2022, GWB § 104 Rn. 36). Dies muss vom Sinn und Zweck der Vorschrift her dahingehend verstanden werden, dass neben dem militärischen Einsatz im rechtlichen Sinne auch der militärische Gebrauch bei einsatzvorbereitenden Maßnahmen, z. B. Übungen sowie Zwecke der Einsatzbereitschaft umfasst sind. **Maßgeblich ist die gewillkürte und objektiv nachprüfbare Zweckbestimmung.** Diese Zweckbestimmung muss objektiv gegeben sein und auch feststehen. Die bloße Behauptung einer militärischen Nutzung und ein nur eventueller Einsatz zu solchen Zwecken reicht nicht aus. Wie auch den Erwägungsgründen zur KI-VO (EG 24) zu entnehmen, kommt es dabei auf den Verwender allein nicht an. So können die Streitkräfte als Teil der Bundeswehr auch zivile bzw. nichtmilitärische Zwecke verfolgen, etwa nach Art. 35 Abs. 2 S. 2 GG beim Einsatz zur Hilfe bei Naturkatastrophen oder einem besonders schweren Unglücksfall. Allerdings sind die in Art. 87a Abs. 3, 4 GG angesprochenen Aufgaben militärischer Natur i.S. des Artikel 2 Abs. 3 UAbs. 2

KI-VO (Burgi/Dreher/Opitz/von Wietersheim, 4. Aufl. 2022, GWB § 104 Rn. 36-40; Martini/Wendehorst, KI-VO 1. Aufl. 2024, KI-VO Art. 2 Rn. 64). Der Einsatz einer KI für die verfassungsrechtlich formulierte Auftragserfüllung der militärischen Verteidigung gemäß Art. 87a Abs. 1 S. 1 GG dürfte ebenfalls grundsätzlich militärischen Zwecken dienen. Die zivile Verteidigung ist von Art. 87a Abs. 1 nicht erfasst (Dreier/Heun, 3. Aufl. 2018, GG Art. 87a Rn. 11).

- c. Neben militärischen sind maßgeblich **Zwecke der Verteidigung** für den GB BMVg relevant.

Unter Verteidigungszwecken können in Anlehnung an Art. 42 EUV alle **militärischen** Maßnahmen zum Schutz des Staatsgebiets vor bewaffneten Angriffen von außen verstanden werden, einschließlich der militärischen Organisation oder der Verteidigung gegen Cyberangriffe anderer Staaten (Martini/Wendehorst/Wendehorst, 1. Aufl. 2024, KI-VO Art. 2 Rn. 64).

Aus EG 24 ergibt sich, dass zum einen nationale Maßnahmen zu Verteidigungszwecken (welche bereits primärrechtlich unter den Kompetenzvorbehalt des Art. 4 Abs. 2 UAbs. 3 EUV fallen und so ohnehin nicht in den Anwendungsbereich der KI-VO fielen) der Bereichsausnahme unterliegen. Dies erfasst auch sämtliche Maßnahmen zu Verteidigungszwecken, die in einem System gegenseitiger kollektiver Sicherheit bzw. im Rahmen von Bündnisverpflichtungen erfolgen. Darüber hinaus sind aber auch alle Maßnahmen, die dem Aufbau einer koordinierten Selbstverteidigung gegenüber bewaffneten Angriffen von außen auf das Unionsgebiet dienen oder mit ihm zusammenhängen, von der Bereichsausnahme erfasst (vgl. Geiger/Khan/Kotzur/Kirchmair/Geiger/Kirchmair, 7. Aufl. 2023, EUV Art. 42 Rn. 7).

Der Tatbestand der Verteidigungszwecke lässt insoweit Interpretationsspielraum, als nicht geklärt ist, inwieweit der in der Bereichsausnahme des Art. 2 Abs. 3 UAbs. 2 KI-VO formulierte Verteidigungsbegriff über den ohnehin bereits im Primärrecht des Art. 4 Abs. 2 UAbs. 3 EUV formulierten Kompetenzvorbehalt hinausgeht. Vom GB BMVg beabsichtigt war ursprünglich eine Ausnahmeregelung, die zumindest über das rein deklaratorische hinausgeht. Gleichzeitig besteht der europarechtliche Grundsatz, dass eine Ausnahmeregelung (die Bereichsausnahme) zu einer bestehenden Ausnahmeregelung (dem Kompetenzvorbehalt), die bereits eng auszulegen ist, ebenfalls eng auszulegen ist. Einmal mehr ist die Begründung im Einzelfall essenziell.

3. Eine Begrenzung der genannten Tatbestände erfolgt weiter durch den Wortlaut „ausschließlich“; danach sind nur solche KI-Anwendungen von der Verordnung ausgenommen, wenn und soweit sie **ausschließlich** für die erwähnten Zwecke eingesetzt werden.

Den Erwägungsgründen (hier: EG 24) ist zu entnehmen, dass ein KI-System, welches zunächst militärischen oder Verteidigungszwecken diene, dann aber für andere Zwecke eingesetzt wird (zivile oder humanitäre, u.a. Strafverfolgungszwecke) ab dem Zeitpunkt unter die Verordnung fällt, ab dem es nicht mehr rein militärischen Zwecken dient. In solchen Fallkonstellationen muss die Einrichtung, welche das KI-System für andere Zwecke als die ausgeschlossenen nutzt, die Konformität mit der Verordnung sicherstellen. Das bedeutet im Ergebnis, dass private Unter-

nehmen, welche solche Systeme, die zuvor der Bereichsausnahme unterfielen, nunmehr für die zivile Nutzung am Markt anbieten wollen, die Vorgaben der KI-VO einhalten müssen. Die Stelle, welche das System (weiterhin) ausschließlich militärisch nutzt, muss dies hingegen nicht.

Umgekehrt fällt nach dem letzten Satz der EG 24 ein KI-System, das zunächst für zivile Zwecke oder Strafverfolgungszwecke in Verkehr gebracht wurde und für Zwecke in den Bereichen Militär, Verteidigung oder nationale Sicherheit verwendet wird, nicht in den Anwendungsbereich der KI-Verordnung, also unter die Bereichsausnahme.

Sogenannte Dual-Use KI-Systeme, die für einen ausgeschlossenen Zweck, nämlich Militär, Verteidigung oder nationale Sicherheit, **und** für einen oder mehrere nicht ausgeschlossene Zwecke, etwa zivile Zwecke oder Strafverfolgungszwecke, in Verkehr gebracht oder in Betrieb genommen werden, fallen in den Anwendungsbereich der Verordnung (vgl. EG 24), wenn sie **(nur vorübergehend oder dauerhaft)** nicht konkret ausschließlich für einen der ausgenommenen Zwecke verwendet werden.

Die Verwendung derartiger KI-Systeme ausschließlich für Zwecke in den Bereichen nationale Sicherheit, Militär und Verteidigung soll aber auch in solchen Fällen weiterhin vom Anwendungsbereich der KI-VO ausgenommen sein und unabhängig davon erfolgen können, dass an anderer Stelle durch eine zivile Einrichtung die Einhaltung der KI-VO sichergestellt werden muss (Rückschluss aus EG 24). Entspricht das KI-System nicht ohnehin bereits der KI-VO, muss derjenige, der die Zweckänderung vornimmt, die Einhaltung der Verordnung sicherstellen (vgl. Martini/Wendehorst/Wendehorst, 1. Aufl. 2024, KI-VO Art. 2 Rn. 65).

Gerade letztgenannte Fallkonstellation macht deutlich, wie schwierig es ist, den Begriff „ausschließlich“ logisch in die Auslegung der Bereichsausnahmeregelung einzubinden. **Gemäß der EG wird ein KI-System also noch „ausschließlich“ für Zwecke z. B. im Bereich Verteidigung verwendet, auch wenn an anderer Stelle eine zivile Einrichtung das gleiche System nutzt.**

4. Die genannten Abgrenzungsfragen machen deutlich, wie wichtig eine gute Begründung im Einzelfall sein wird. Neben dem vielfach erwähnten EG 24 sollten unbedingt auch der **Zweck und die Rechtfertigung der Bereichsausnahme** zur Begründung herangezogen werden.

Durch die Ausnahme des Art. 2 Abs. 3 UAbs. 2 KI-VO soll verhindert werden, „dass die Verteidigungsfähigkeit und Sicherheit der Mitgliedstaaten der Union unterminiert wird. In militärischen Konfliktfällen, bei der Abwehr hybrider Angriffe und ganz allgemein im Rahmen der nachrichtendienstlichen Aufklärungsarbeit lassen sich Vorgaben wie die der DS-GVO unter der KI-VO oftmals nicht sinnvoll einhalten. Darüber hinaus würden die in der KI-VO vorgesehenen Transparenzvorschriften hinsichtlich der Funktionsweise der KI-Systeme verteidigungspolitischen Interessen offenkundig zuwiderlaufen.“ (Martini/Wendehorst/Wendehorst, 1. Aufl. 2024, KI-VO Art. 2 Rn. 54)

Berücksichtigt werden sollen die Besonderheiten der Verteidigungspolitik der Mitgliedstaaten und der in Titel V Kapitel 2 EUV abgedeckten gemeinsamen Verteidigungspolitik der Union, die dem Völkerrecht unterliegen. Dieses bildet stattdessen den grundlegenden Rechtsrahmen für die Regulierung von KI-Systemen im Zusammenhang mit der Anwendung tödlicher Gewalt und sonstigen KI-Systemen im Zusammenhang mit Militär- oder Verteidigungstätigkeiten.

5. Hinzuweisen ist weiterhin auf **UAbs. 3** des Art. 2 Abs. 3 KI-VO. Dieser könnte manche Fälle noch unter die Bereichsausnahme fassen, welche nicht unter UAbs. 2 zu subsumieren sind.

Die KI-VO gilt danach entgegen der allgemeinen Regel in Abs. 1 lit. c unabhängig von ihrer allgemeinen Zweckbestimmung nicht für KI-Systeme, die in Drittstaaten in Verkehr gebracht oder in Betrieb genommen werden, wenn der Output zwar in der Union verwendet wird, dies aber ausschließlich für militärische, verteidigungspolitische oder die nationale Sicherheit betreffende Zwecke erfolgt. Durch diese Regelung soll die militärische und nachrichtendienstliche Zusammenarbeit der Union und ihre Mitgliedstaaten mit Drittstaaten abgesichert werden. Zwar wird in den meisten Fällen bereits UAbs. 2 des Abs. 3 eingreifen, doch dient UAbs. 3 der Lückenfüllung für Situationen, in denen die verwendeten KI-Systeme auch für andere Zwecke verwendet werden (z. B. für Strafverfolgungszwecke). (Martini/Wendehorst/Wendehorst, 1. Aufl. 2024, KI-VO Art. 2 Rn. 55)

V. Text des dazugehörigen interpretatorischen Erwägungsgrundes (EG) 24:

Wenn und soweit KI-Systeme mit oder ohne Änderungen für Zwecke in den Bereichen Militär, Verteidigung oder nationale Sicherheit in Verkehr gebracht, in Betrieb genommen oder verwendet werden, sollten sie vom Anwendungsbereich dieser Verordnung ausgenommen werden, unabhängig von der Art der Einrichtung, die diese Tätigkeiten ausübt, etwa ob es sich um eine öffentliche oder private Einrichtung handelt. In Bezug auf die Zwecke in den Bereichen Militär und Verteidigung gründet sich die Ausnahme sowohl auf Artikel 4 Absatz 2 EUV als auch auf die Besonderheiten der Verteidigungspolitik der Mitgliedstaaten und der in Titel V Kapitel 2 EUV abgedeckten gemeinsamen Verteidigungspolitik der Union, die dem Völkerrecht unterliegen, was daher den geeigneteren Rechtsrahmen für die Regulierung von KI-Systemen im Zusammenhang mit der Anwendung tödlicher Gewalt und sonstigen KI-Systemen im Zusammenhang mit Militär- oder Verteidigungstätigkeiten darstellt. In Bezug auf die Zwecke im Bereich nationale Sicherheit gründet sich die Ausnahme sowohl auf die Tatsache, dass die nationale Sicherheit gemäß Artikel 4 Absatz 2 EUV weiterhin in die alleinige Verantwortung der Mitgliedstaaten fällt, als auch auf die besondere Art und die operativen Bedürfnisse der Tätigkeiten im Bereich der nationalen Sicherheit und der spezifischen nationalen Vorschriften für diese Tätigkeiten. Wird ein KI-System, das für Zwecke in den Bereichen Militär, Verteidigung oder nationale Sicherheit entwickelt, in Verkehr gebracht, in Betrieb genommen oder verwendet wird, jedoch vorübergehend oder ständig für andere Zwecke verwendet, etwa für zivile oder humanitäre Zwecke oder für Zwecke der Strafverfolgung oder öffentlichen Sicherheit, so würde dieses System in den Anwendungsbereich

dieser Verordnung fallen. In diesem Fall sollte die Einrichtung, die das KI-System für andere Zwecke als Zwecke in den Bereichen Militär, Verteidigung oder nationale Sicherheit verwendet, die Konformität des KI-Systems mit dieser Verordnung sicherstellen, es sei denn, das System entspricht bereits dieser Verordnung. KI-Systeme, die für einen ausgeschlossenen Zweck, nämlich Militär, Verteidigung oder nationale Sicherheit, und für einen oder mehrere nicht ausgeschlossene Zwecke, etwa zivile Zwecke oder Strafverfolgungszwecke, in Verkehr gebracht oder in Betrieb genommen werden, fallen in den Anwendungsbereich dieser Verordnung, und Anbieter dieser Systeme sollten die Einhaltung dieser Verordnung sicherstellen. In diesen Fällen sollte sich die Tatsache, dass ein KI-System in den Anwendungsbereich dieser Verordnung fällt, nicht darauf auswirken, dass Einrichtungen, die Tätigkeiten in den Bereichen nationale Sicherheit, Verteidigung oder Militär ausüben, KI-Systeme für Zwecke in den Bereichen nationale Sicherheit, Militär und Verteidigung verwenden können, unabhängig von der Art der Einrichtung, die diese Tätigkeiten ausübt, wobei die Verwendung vom Anwendungsbereich dieser Verordnung ausgenommen ist. Ein KI-System, das für zivile Zwecke oder Strafverfolgungszwecke in Verkehr gebracht wurde und mit oder ohne Änderungen für Zwecke in den Bereichen Militär, Verteidigung oder nationale Sicherheit verwendet wird, sollte nicht in den Anwendungsbereich dieser Verordnung fallen, unabhängig von der Art der Einrichtung, die diese Tätigkeiten ausübt.

5.2 Erläuterung zur Einführung eines Risikomanagements Künstlicher Intelligenz im Geschäftsbereich des BMVg

5.2.1 Risikoidentifizierung, -bewertung und -steuerung

1. Wie in dem Abschnitt 3.3.1 dargestellt, muss ein RM KI nach Vorgaben der EU KI-VO und NATO Blueprint etabliert werden, um Schwachstellen, Bedrohungen und Risiken frühzeitig zu identifizieren, zu beheben oder zu managen.
2. Das RM KI des GB BMVg muss – wenn die EU KI VO in dem entsprechenden Bereich gilt (siehe Abb. 6) – bei Hochrisikosystemen u. a. eine Risikofolgeabschätzung für Grundrechte und über den gesamten Lebenszyklus von KI-Systemen die Ermittlung, Analyse, Bewertung und Mitigation von Risiken durch konkrete Maßnahmen ermöglichen.⁴⁸ Insgesamt gilt es eine vorgabenkonforme und möglichst aufwandsarme Umsetzung für den GB BMVg zu wählen und in einer nachfolgenden AR „Risikomanagement KI für den GB BMVg“ zu spezifizieren.

5.2.1.1 KI-bezogene Risiken im Geschäftsbereich des BMVg

3. Um ein erfolgreiches RM KI zu betreiben, werden mögliche Risiken für KI-Systeme im Kontext GB BMVg gesammelt und – dem Beispiel des NIST-Risk-Management Framework for AI folgend⁴⁹ – in Risikokategorien geclustert.⁵⁰ Diese sind für den GB BMVg „Datenrisiken“, „Modellrisiken“, „Operationelle Risiken“, „Anwenderinnen- und Anwenderrisiken“ und „Regulatorik“.
4. Die Definitionen der einzelnen Risiken sind nachfolgend (siehe Anlage 5.2.2) zu finden. Diese Liste an Risiken hat jedoch keinen Anspruch auf Vollständigkeit, da die Risiken sehr von den individuellen KI-Systemen abhängig sind. Die Liste gibt daher eine erste Orientierung und zeigt auf, welche möglichen Risiken für den GB BMVg existieren und über die verschiedenen KI-Lebenszyklusphasen auftreten können.

⁴⁸ EU KI-VO, Art. 9.

⁴⁹ Das NIST Risk-Management Framework for AI wird sowohl in der EU KI-VO, als auch in dem NATO Blueprint on Responsible AI als Referenzrahmenwerk für das KI-RM erwähnt.

⁵⁰ NIST: AI Risk Management Framework.

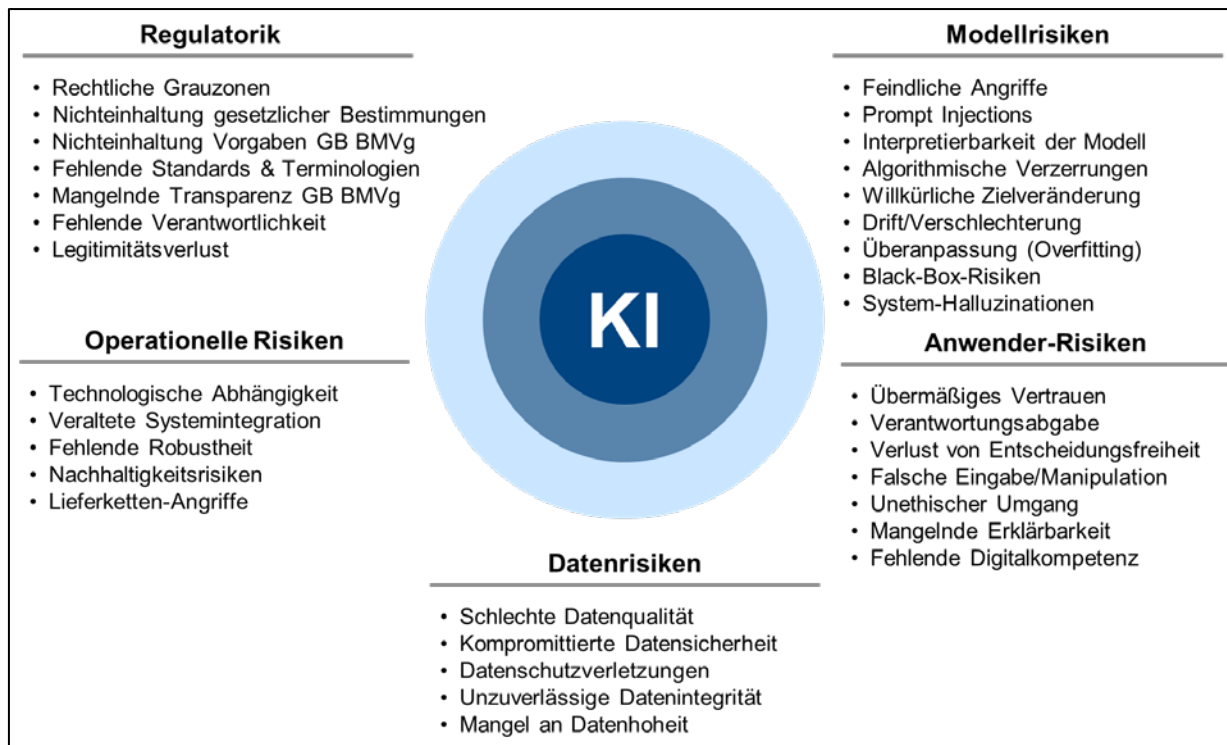


Abb. 2: Übersicht KI-Risiken (Risk Map).

5.2.1.2 Bewertung der KI-bezogenen Risiken für den Geschäftsbereich des BMVg

5. Da nur eine begrenzte Menge an Ressourcen im GB für das RM von KI-Systemen zur Verfügung steht, können nicht alle KI-bezogenen Risiken im gleichen Ausmaß gesteuert werden. Dies macht eine Priorisierung notwendig und gilt sowohl für die übergreifende Perspektive GB BMVg als auch für die einzelnen KI-Systeme.

6. Das Vorgehen bei der Priorisierung wird sich auf ein durch die NATO beschriebenes Vorgehen (z. B. die Erstellung eines „Risiko-Profiles“) und die zugehörigen Werkzeuge, wie die „NATO ACO Severity Matrix“, stützen.

5.2.1.3 Steuerung der KI-bezogenen Risiken des Geschäftsbereichs des BMVg

7. Um die priorisierten KI-bezogenen Risiken (Risiko-Profil KI) zielgerichtet zu steuern, sind entsprechende Strukturen und Abläufe für das RM KI notwendig. Oberstes Ziel ist es, das RM KI in den bestehenden Strukturen und Abläufen des GB BMVg zu betreiben (z. B. in den RM-Strukturen für herkömmliche IT-Systeme).

8. Die Notwendigkeit der Anpassung bestehender und ggf. Schaffung neuer Abläufe und Daten- bzw. Informationsstrukturen ist aufgrund der regulatorischen Vorgaben für KI-Systeme und das RM-KI aber nicht auszuschließen.

9. Die Anpassung bestehender oder Schaffung neuer Strukturen und Abläufe ist abhängig von den neu auszuführenden Aufgaben. Diese (KI-spezifischen) Aufgaben im RM sind u. a.

- Analyse des Risikos eines KI-Systems über dessen gesamten Lebenszyklus (z. B. mit Zuhilfenahme des Risiko-Profiles), inklusive Grundrechte-Folgeabschätzung vor der Einführung,
- kontinuierliche Entwicklung und Koordinierung von notwendigen Mitigationsmaßnahmen (z. B. abgeleitet aus den jeweiligen KI Risiko-Profilen),
- Koordinieren von notwendigen Feedbackschleifen⁵¹ zwischen Anwendenden, Beschaffenden, Entwicklenden und Anbietenden,
- zentrale Steuerung und Beratung zur Mitigation von auftretenden schwerwiegenden unbeabsichtigten Effekten beim Einsatz von KI-Systemen bzw. KI-spezifischen Problemen (Meldesystem),
- laufende, automatisierte Risikoüberwachung der Systemleistung (z. B. möglicher Schwachstellen und unvorhergesehener Auswirkungen)⁵²,
- Durchführen regelmäßiger Audits nach einer noch zu entwickelnden Checkliste, damit die KI-Systeme auch im laufenden Betrieb den ethischen, rechtlichen und technischen Standards entsprechen,
- revisionssicheres Archivieren wesentlicher Aktivitäten und Anpassungen im Einsatz an KI-Systemen zur Rückverfolgbarkeit und Nachvollziehbarkeit der Ergebnisse und
- Kontrolle des Zugriffs durch Angehörige dritter Staaten.

10. **Abläufe:** Die Abläufe sind von den bestehenden oder neu zu schaffenden Daten- bzw. Informationsstrukturen abhängig. Je nachdem welche Akteure involviert sind, müssen die entsprechenden Abläufe angepasst oder neu geschaffen werden.

5.2.2 Risiko-Definitionen

11. In diesem Abschnitt werden Definitionen für Beispiele von Risiken dargestellt, die insbesondere oder ausschließlich bei der Nutzung von KI auftreten können. Diese Beispiele haben keinen Anspruch auf Vollständigkeit.

5.2.2.1 Regulatorik

- **Rechtliche Grauzonen:** Aktuell bewegen sich viele KI-Systeme rechtlich noch in Graubereichen. Das birgt Risiken, da sich Rechtsauffassungen und Urteile erst noch entwickeln und kontinuierlich beobachtet werden müssen.

⁵¹ KI-Systeme sollten daher grundsätzlich nach der Auftragslogik und nicht nach einer Befehlslogik eingesetzt werden: Sobald das KI-System beginnt zu arbeiten, kann im weiteren Verlauf durch die ersten Rückmeldungen und Ergebnisse des KI-Systems entschieden werden, wie weiter vorgegangen wird (Feedbackschleifen).

⁵² Kennzahlen wie „Genauigkeit“, „Zuverlässigkeit“ und „Fehlerraten“ spielen eine zentrale Rolle, um sicherzustellen, dass das System den festgelegten Qualitätsstandards entspricht.

- **Nichteinhaltung von gesetzlichen Bestimmungen:** Es kann zu rechtlichen Sanktionen führen, wenn gesetzliche Vorschriften wie die DSGVO oder branchenspezifische Richtlinien nicht eingehalten werden.
- **Nichteinhaltung von Vorgaben:** Spezifische Vorgaben des GB werden von den Nutzenden der KI-Systeme nicht beachtet.
- **Fehlende Standards und Terminologien:** Das Fehlen einheitlicher Standards und Terminologien für KI-Systeme führt u. a. zu einer heterogenen IT-Landschaft, unterschiedlichen Qualitätsniveaus und Missverständnissen.
- **Mangelnde Transparenz:** Der GB riskiert einen Vertrauensverlust innerhalb der Organisation und in der Öffentlichkeit, wenn KI-Systeme nicht transparent und verantwortungsvoll eingesetzt werden.
- **Fehlende Verantwortlichkeiten:** Viele Organisationen verfügen noch nicht über geeignete Governance-Strukturen für den Einsatz von KI. Dadurch ist eine umfassende Überwachung der Systeme oft nicht gewährleistet.
- **Legitimitätsverlust:** KI-Systeme können öffentlich auf Ablehnung stoßen und die Legitimität untergraben, selbst wenn sie technisch korrekt arbeiten.

5.2.2.2 Datenrisiken

- **Kompromittierte Datensicherheit:** Angriffe können dem GB ernsthafte Probleme bereiten, indem sie Datensätze manipulieren, denen KI-Technologien zugrunde liegen, z. B. durch unbefugten Zugriff, Datenverlust und Verletzung der Vertraulichkeit.
- **Schlechte Datenqualität:** Die Qualität der Daten ist unzureichend oder verschlechtert sich im Zeitverlauf.
- **Datenschutzverletzungen:** KI-Systeme verarbeiten häufig sensible personenbezogene Daten, die anfällig für Datenschutzverletzungen sein können.
- **Unzuverlässige Datenintegrität:** KI-Modelle sind nur so zuverlässig wie ihre Trainingsdaten. Verfälschte, verzerrte oder doppelte Datensätze können zu ungenauen Ergebnissen und fehlerhaften Entscheidungen führen.
- **Mangel an Datenhoheit:** Bei der Zusammenarbeit mit externen Dienstleistern: z. B. im Rüstungsbereich, besteht das Risiko, den Zugriff auf kritische Daten zu verlieren oder die Kontrolle darüber abzugeben, dies kann grundsätzlich auch bei der Weitergabe an Drittstaaten der Fall sein.

5.2.2.3 Modellrisiken

- **Feindliche Angriffe:** Angriffe manipulieren Eingabedaten, um KI-Systeme gezielt zu falschen Vorhersagen oder Entscheidungen zu verleiten.
- **Prompt Injections:** Angriffe auf große Sprachmodelle, bei denen manipulierte Eingaben genutzt werden, um Schutzmechanismen zu umgehen oder sensible Informationen zu extrahieren.

- **Interpretierbarkeit der Modelle:** Komplexe Modelle sind schwer nachvollziehbar. Dies erschwert das Verständnis von Entscheidungen, behindert die Aufdeckung von Verzerrungen und schwächt das Vertrauen.
- **Algorithmische Verzerrungen:** KI-Algorithmen können Vorurteile (statistische Verteilungen) aus Trainingsdaten übernehmen, was zu potenziell diskriminierenden Ergebnissen führen kann.
- **Willkürliche Zielveränderung:** Das KI-System verändert unbemerkt oder unbeabsichtigt seine Zielausrichtung.
- **Drift/Verschlechterung:** Modell-Drift beschreibt die Leistungsverschlechterung eines KI-Systems durch veränderte Daten oder Datenbeziehungen.
- **Überanpassung (Overfitting):** Ein überangepasstes Modell liefert zwar gute Ergebnisse im Training, scheitert jedoch in der praktischen Anwendung aufgrund mangelnder Generalisierbarkeit.
- **Black-Box-Risiken:** Die fehlende Transparenz und Nachvollziehbarkeit von KI-Entscheidungen verursacht rechtliche und ethische Probleme bzgl. Verantwortung und Kontrolle.
- **System-Halluzinationen:** Bei unzureichenden oder schlechten Daten kann die KI falsche oder irreführende Ergebnisse hervorbringen, da das KI-System potenzielle Muster erkennen und vervollständigen kann, was zu unrichtigen Aussagen führen kann.

5.2.2.4 Anwenderinnen- und Anwender-Risiken

- **Übermäßiges Vertrauen:** Nutzende müssen sowohl den Nutzen als auch die Grenzen von KI-gestützten Entscheidungshilfen verstehen.
- **Verantwortungsabgabe/Verantwortungsdiffusion:** Trotz KI-Unterstützung liegt die Entscheidungsverantwortung weiterhin beim Menschen. Die Rolle der KI ist rein unterstützend und sollte stets auf Plausibilität geprüft werden.
- **Deskilling:** Das Risiko durch ständige Nutzung von KI, die Fähigkeiten von Personal verloren gehen, welche nicht mehr bei einem Ausfall der KI kompensiert werden können.
- **Verlust von Entscheidungsfreiheit:** KI-Systeme dürfen die Entscheidungsfindung nicht durch manipulative Ausgabeformate oder suggestive Formulierungen beeinflussen; die schädliche Gewöhnung an KI-generierte Entscheidungsvorschläge ist durch entsprechendes Training zu vermeiden.
- **Falsche Eingabe/Manipulation:** KI-Systeme müssen robust gegenüber fehlerhaften oder absichtlich manipulierten Eingaben sein – auch durch Mitarbeitende.
- **Unethischer Umgang:** KI-Systeme könnten missbraucht werden, um Inhalte zu erzeugen oder zu verbreiten, die den ethischen Grundsätzen des GB widersprechen.
- **Herausfordernde Erklärbarkeit:** Erklärbarkeit in der KI bezieht sich auf die Fähigkeit, die von KI-Systemen generierten Ergebnisse zu verstehen und zu begründen. Mangelnde Erklärbarkeit kann das Vertrauen in die Person und Organisation untergraben und zu Reputationsschäden führen.

- **Fehlende Digitalkompetenz:** Ein sicherer und verantwortungsvoller Umgang mit KI-Systemen erfordert angemessene Schulungen und digitale Kompetenzen.

5.2.2.5 Operationelle Risiken

- **Technologische Abhängigkeit von Anbietern:** Ein hoher Grad an Abhängigkeit schränkt die strategische Flexibilität ein und kann langfristig zu steigenden Kosten führen.
- **Veraltete Systemintegration:** Die Einbindung moderner KI-Systeme in bestehende IT-Landschaften kann durch veraltete Strukturen erheblich erschwert werden.
- **Fehlende Robustheit:** KI-Systeme müssen gegen äußere Einflüsse wie Stromausfälle, Netzwerkausfälle oder Systemkomponenten-Fehler widerstandsfähig sein.
- **Nachhaltigkeitsrisiken:** Vernachlässigte Wartung, hoher Energieverbrauch oder ineffiziente Ressourcennutzung können die langfristige Funktionsfähigkeit gefährden.
- **Angriffe auf die Lieferkette:** Feindliche Akteure könnten Schwachstellen bei Drittanbietern oder während der Entwicklung, Integration oder Wartung ausnutzen.
- **Rechtslage in anderen Staaten:** Unternehmen oder Angehörige aus bestimmten Staaten könnten zur Herausgabe von Informationen rechtlich verpflichtet werden.

5.3 Abkürzungsverzeichnis

Abkürzung	Bedeutung
AA	Auswärtiges Amt
AI	Artificial Intelligence
AIN	Ausrüstung, Informationstechnik und Nutzung (OrgBer)
AR	Allgemeine Regelung
A&S	Analytics & Simulation
AuQITGBBMVg	Ausbildung und Qualifizierung im Bereich der IT im GB BMVg
AusbZ CIR	Ausbildungszentrum Cyber- und Informationsraum
BAMAD	Bundesamt für den Militärischen Abschirmdienst
BDSG	Bundesdatenschutzgesetz
BfDI	Bundesbeauftragte bzw. Bundesbeauftragter für Datenschutz und Informationsfreiheit
BMVg	Bundesministerium der Verteidigung
Bw	Bundeswehr
CAIO	Chief Artificial Intelligence Officer
CCW	Certain Conventional Weapons
CDO	Chief Data Officer
CIO	Chief Information Officer
CIR	Cyber- und Informationsraum
DEK	Datenethikkommission der Bundesregierung
DEU	Deutschland/deutsch
DSGVO	Datenschutz-Grundverordnung
DSt	Dienststelle
EDA	European Defence Agency
EDT	Emerging Disruptive Technologies
EG	Erwägungsgrund
EGovG	E-Government-Gesetz
EU	Europäische Union
EU KI-VO	EU KI-Verordnung
EuGH	Europäischer Gerichtshof
EUV	Vertrag über die Europäische Union
FF	Federführend
FMN	Framework Federated Mission Networking
FRA	Frankreich
F&T	Forschung und Technologie
GB	Geschäftsbereich

Abkürzung	Bedeutung
GBR	Vereinigtes Königreich
GeNu	Genehmigung zur Nutzung
IAT	Integriertes Aufgabenteam
IC	Abteilung Innovation & Cyber
ICT	Integriertes Clusterteam
ICU	Infrastructure, Cloud Base and User Equipment
InnoZBw	Innovationszentrum der Bundeswehr
InspCIR	Inspekteur bzw. Inspekteurin Cyber- und Informationsraum
IPT	Integriertes Projektteam
ISO	International Organization for Standardization
KD	Konzeptionelles Dokument
KdoCIR	Kommando Cyber- und Informationsraum
KI	Künstliche Intelligenz
LAWS	Letales autonomes Waffensystem
LV/BV	Landes- und Bündnisverteidigung
MDO	Multi-Domain Operation
MilOrgBer	Militärischer Organisationsbereich
ML	Machine Learning
NATO	North Atlantic Treaty Organization
ngBer	nachgeordneter Bereich
NIST	National Institute of Standards and Technology
OrgBer	Organisationsbereich
pCloudBw	private Cloud Bw
PBN	Projektbezogene Bedarfsdeckung und Nutzung
PlgABw	Planungsamt der Bundeswehr
POC	Point of Contact
PRU	NATO Principles of Responsible Use of Artificial Intelligence in Defence
RAI	NATO Responsible AI
REAIM	Responsible AI in the Military Domain
RM	Risikomanagement
RO	Recht und Organisation
SD	Strategisch-politisches Dokument
SDD	Software Defined Defence
StVH	Stabs- und Verwaltungshandeln
SVI-Strategie	Sicherheits- und Verteidigungs-Industriestrategie
TSK	Teilstreitkräfte

Abkürzung	Bedeutung
UAbs.	Unterabsatz
UNIDIR	United Nations Institute for Disarmament Research
USA	Vereinigte Staaten von Amerika
VLBw	Virtuelle Lernumgebung der Bundeswehr
VN	Vereinte Nationen
V&V	Verifikation und Validierung
WaSys	Waffensystem
ZDigBw	Zentrum Digitalisierung der Bundeswehr und Fähigkeitsentwicklung Cyber- und Informationsraum

