AI Assurance: Vertrauen, Sicherheit und schnelle Einsatzfähigkeit für KI-Systeme in der Bundeswehr

Kim-Laura Wöhlk, IABG; Rafal Kaluga, IABG, Bastian Bernhardt, IABG

Kim-Laura Wöhlk

Foto: IABG

Mit der fortschreitenden Digitalisierung und der Einführung disruptiver Technologien steht die Bundeswehr vor einer doppelten Zeitenwende. Neue technologische Möglichkeiten revolutionieren militärische Operationen und Entscheidungsprozesse, während zugleich die wachsende Komplexität und Vernetzung dieser Systeme neue Anforderungen an deren Sicherheit und Verlässlichkeit stellen. Besonders die Integration von künstlicher Intelligenz (KI) in sicherheitskritische Anwendungen erfordert ein hohes Maß an Vertrauen. Um dies zu gewährleisten, rückt das Konzept der **AI Assurance**, der systematischen Absicherung von KI-Systemen, in den Mittelpunkt moderner Absicherungsstrategien.

Rafal Kulaga

Foto: IABG

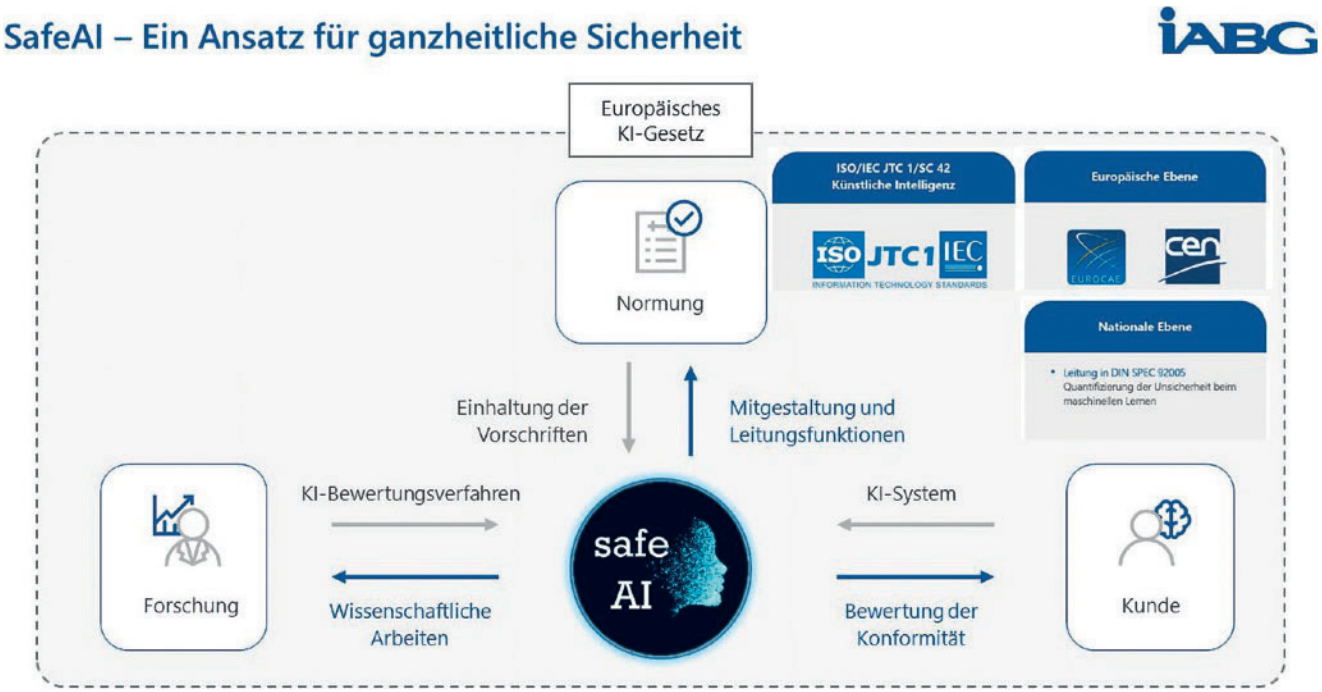
Die Einsatzmöglichkeiten von KI in der Bundeswehr sind vielfältig und reichen von der Automatisierung logistischer Abläufe über die Verarbeitung großer Datenmengen aus Sensor- und Aufklärungssystemen bis hin zur Entscheidungsunterstützung in dynamischen, sich schnell verändernden Szenarien. Die Vorteile liegen auf der Hand: KI setzt an, wo klassische Lösungen an ihre Grenzen kommen. Sie kann enorme Datenmengen analysieren, Muster erkennen und auf dieser Basis Handlungsempfehlungen ableiten, die weit über die Fähigkeiten traditioneller Systeme hinausgehen. Doch so beeindruckend diese Fortschritte auch sind, sie bringen Herausforderungen mit sich, die nicht ignoriert werden dürfen. KI-Systeme müssen nicht nur zuverlässig funktionieren („Konformitätserklärung“), sondern auch robust gegen Angriffe sein, nachvollziehbare Entscheidungen treffen und regulatorischen sowie ethischen Anforderungen genügen. Ein besonders kritischer Aspekt ist die Frage der Verlässlichkeit und Robustheit. KI-Systeme sind komplex und basieren oft auf maschinellen Lernverfahren, die große Mengen an Daten nutzen, um Modelle zu trainieren. Diese Modelle sind jedoch nicht immun gegen Manipulationen oder unvorhergesehene Einflüsse. Schon kleine Änderungen an den Eingangsdaten können die Funktionsweise von KI-Systemen erheblich beeinträchtigen und in sicherheitskritischen Anwendungen fatale Folgen haben. Dies macht es unabdingbar, dass KI nicht nur leistungsfähig, sondern auch resilient gegen äußere Eingriffe gestaltet wird.

safeAI: Ein Ansatz für ganzheitliche Sicherheit

Bastian Bernhardt

Foto: IABG

Moderne Absicherungskonzepte wie der safeAI-Ansatz der IABG greifen auf einen ganzheitlichen Ansatz zurück, um diese Herausforderungen zu bewältigen. Neben funktionaler Sicherheit legt safeAI den Fokus auf die Gestaltung von sicheren und vertrauenswürdigen KI-Systemen. Dies erfordert eine Kombination technischer und organisatorischer Maßnahmen, die auf gesamten Lebenszyklus eines KI-Systems wirken. safeAI setzt auf fortschrittliche Test- und Validierungsumgebungen, die realistische Einsatzszenarien nachbilden und Schwachstellen frühzeitig aufdecken. Dabei wird nicht nur die technische Leistung geprüft, sondern auch die Fähigkeit der KI, sich in unerwarteten oder dynamischen Situationen robust zu verhalten. Ergänzend integriert safeAI statistisch fundierte Sicherheitsmaßnahmen, wie die Bewertung der Unsicherheit oder der Robustheit gegenüber Angriffen. Diese Maßnahmen gewährleisten, dass die Systeme unter realen Bedingungen zuverlässig und sicher arbeiten. Ein weiterer Schwerpunkt von safeAI liegt auf der Erklärbarkeit und Transparenz. Gerade in sicherheitskritischen Kontexten müssen die Entscheidungen einer KI für die Nutzer nachvollziehbar sein. safeAI-Ansätze setzen hier auf Methoden der erklärbaren KI (Explainable AI, XAI), die Entscheidungswege offenlegen und so das Vertrauen der Anwender stärken. Für die Bundeswehr, die in hochkomplexen und oft unvorhersehbaren Szenarien agiert, ist diese Transparenz essenziell, um fundierte Entscheidungen treffen zu können.



safeAI – Ein Ansatz für ganzheitliche Sicherheit

Foto: IABG

Regulatorische und ethische Rahmenbedingungen

Neben der technischen Absicherung steht safeAI auch für die Einhaltung regulatorischer und ethischer Anforderungen. KI-Systeme müssen nicht nur zuverlässig funktionieren, sondern auch den strengen Vorgaben nationaler und internationaler Standards genügen. Mit safeAI unterstützt und entwickelt die IABG zertifizierbare Prozesse, begleitet Standardisierungsverfahren und passt das „Absicherungskonzept“ an. Darüber hinaus trägt die IABG mit safeAI aktiv führend zur Entwicklung von KI-Standards bei, etwa durch die Initiierung der DIN SPEC 92005 und die Mitwirkung an der ISO/IEC AWI TS 25223.

Die Bedeutung eines ganzheitlichen Ansatzes

Ein moderner Ansatz wie safeAI zeigt, dass AI Assurance nicht nur technische Innovation bedeutet, sondern auch ein ganzheitliches Denken erfordert. Die Kombination aus Sicherheitsanalysen, leistungsfähigen Simulationsumgebungen und einer systematischen Validierung trägt dazu bei, Vertrauen in KI-Systeme zu schaffen. safeAI betont dabei die Rolle einer nachhaltigen Sicherheitskultur, die technische Exzellenz mit organisatorischer Verantwortung verbindet.

Fazit: Vertrauen als Basis für Innovation

AI Assurance, gestützt durch innovative Ansätze wie safeAI, bildet die Grundlage, um Chancen für die Bundeswehr sicher und verantwortungsvoll zu erschließen. Dabei verbindet safeAI technische, regulatorische und ethische Aspekte und liefert somit einen entscheidenden Beitrag für die Einsatzfähigkeit und Resilienz moderner Streitkräfte.