BSI

Kriterien für sicheren KI-Einsatz

[26.06.2025] Beim Zukunftskongress hat das BSI einen Kriterienkatalog für den sicheren Einsatz generativer KI in der Bundesverwaltung vorgestellt. Die Publikation adressiert KI-typische Sicherheitsrisiken und soll als Orientierungshilfe für Behörden dienen.

Im Rahmen des 11. Zukunftskongresses Staat & Verwaltung (23. bis 25. Juni 2025, Berlin) hat das Bundesamt für Sicherheit in der Informationstechnik (BSI) den neuen Kriterienkatalog zum Einsatz von generativer Künstlicher Intelligenz (KI) in der Bundesverwaltung präsentiert. Aufgrund ihrer vielseitigen Einsatzmöglichkeiten in der Generierung, Modifikation und Analyse verschiedener Inhalte gewinnen generative KI-Modelle seit Ende 2022 rasant an Bedeutung und werden als Chance für die Digitalisierung der öffentlichen Verwaltung begriffen. Einsatzszenarien für insbesondere textverarbeitende KI-Modelle liegen beispielsweise im Bereich von Chatbots und der Textzusammenfassung, -analyse sowie -übersetzung. Immer häufiger werden Funktionen generativer KI-Modelle auch in eigene Anwendungen der Bundesverwaltung integriert. Gleichzeitig kann die IT-Sicherheit von generativen KI-Anwendungen, im Vergleich zu Standard-IT-Komponenten, über neuartige Angriffsvektoren beeinträchtigt werden. So können so genannte (Indirect) Prompt Injections die Ausgaben von Chatbots zielgerichtet manipulieren. Die Konsequenzen hieraus können die Verbreitung von Schadsoftware oder das Abfließen sensitiver Daten sein.

Erst Kriterienkatalog, dann Mindeststandard

Um diesen neuen Angriffsvektoren adäquat zu begegnen, hat das BSI einen Kriterienkatalog entwickelt. Dieser definiert Anforderungen, die bei der Integration extern bereitgestellter generativer KI-Modelle in eigene Anwendungen der Bundesverwaltung erfüllt werden sollten, um ein Mindestsicherheitsniveau zu erreichen. Der Katalog richtet sich an IT-Sicherheitsbeauftragte oder vergleichbare Stellen. Zusammen mit den im März 2025 vorgelegten „Leitlinien für den Einsatz Künstlicher Intelligenz in der Bundesverwaltung“ will die Publikation Behörden eine Orientierungshilfe bieten. „Die Einsatzszenarien für KI werden immer mannigfaltiger - die Verwaltung braucht jetzt Orientierung, um frühzeitig die Chancen von KI zu ergreifen und Anwendungen sicher einsetzen zu können. Das BSI bietet der Verwaltung mit dem Kriterienkatalog diese Orientierung und eine konkrete erste Hilfestellung, um generative KI sicher zu nutzen“, so BSI-Vizepräsident Thomas Caspers. Perspektivisch will das BSI zudem einen verbindlichen Mindeststandard zum Einsatz von generativer KI in der Bundesverwaltung veröffentlichen. Daher hat die Sicherheitsbehörde nach eigenem Bekunden bereits jetzt hohes Interesse an Rückmeldungen zum Kriterienkatalog (E-Mail: ki-kontakt@bsi.bund.de).

(sib)

• Kriterienkatalog des BSI zur Integration von extern bereitgestellten generativen KI-Modellen in eigene Anwendungen
• Leitlinien für den Einsatz Künstlicher Intelligenz in der Bundesverwaltung

Stichwörter: Künstliche Intelligenz, BSI, IT-Sicherheit