TÜV Rheinland: KI-Systeme effektiv absichern

Neue Anforderungen an IT-Sicherheit durch generative KI und Large Language Models / Bedrohungen erkennen, Schwachstellen schließen – Penetration Testing als wirksame Maßnahme / Neues Whitepaper mit Handlungsempfehlungen für Unternehmen / www.tuv.com/pentest

Köln, 4. November 2025. Generative Künstliche Intelligenz (KI) revolutioniert nicht nur zahlreiche Branchen, sie schafft auch neue Angriffsflächen und sicherheitstechnische Herausforderungen – etwa, wenn Unternehmen große Sprachmodelle in Chatbots, Assistenzsystemen oder automatisierten Entscheidungsprozessen einsetzen. Doch welche spezifischen Risiken entstehen bei der Nutzung von Large Language Models (LLMs)? Darauf geht TÜV Rheinland im aktuellen Whitepaper „Ist Ihr KI-System sicher?“ ein. Zudem zeigen die Cybersecurity-Fachleute, wie Unternehmen ihre KI-Anwendungen effektiv absichern können.

Angriffe durch manipulierte Eingaben und Trainingsdaten möglich

Das Whitepaper beschreibt, wie KI-Systeme angegriffen werden können. Ein Beispiel sind sogenannte Prompt Injections, bei denen Angreifer mit ihrer Eingabe das Modell manipulieren, damit es sich unvorhersehbar verhält oder Informationen preisgibt – und zwar solche, die nicht zugänglich sein sollten. Weitere Risiken sind der unsichere Umgang mit den Ergebnissen der generativen KI – etwa, indem Nutzer nicht validierte Codes ausführen – und die Manipulation von Trainingsdaten durch einen Angreifer.

Sowohl Angriffe als auch der falsche Umgang mit KI-Ergebnissen können fatale Folgen haben: von Datenlecks über fehlerhafte Entscheidungen bis hin zu wirtschaftlichen Schäden. Daher ist ein systematisches Risikomanagement für Unternehmen unerlässlich – nicht zuletzt fordern dies auch zunehmend Regulierungen wie der EU AI Act. „Unternehmen müssen ihre Sicherheitskonzepte anpassen, um den Risiken von KI-Systemen gerecht zu werden“, erklärt Daniel Hanke, Experte für KI-Sicherheit bei TÜV Rheinland.

Penetration Testing als Schlüssel zur KI-Sicherheit

Eine der wirksamsten Maßnahmen, um Bedrohungen frühzeitig zu erkennen und Schwachstellen zu schließen, ist das Penetration Testing (Pentest): Im kontrollierten Rahmen simulieren Fachleute dabei Angriffe auf KI-Systeme, um

potenzielle Schwachstellen zu identifizieren und zu beheben. Methoden wie Black-Box- und Gray-Box-Tests werden dabei an die Anforderungen von generativer KI angepasst. „KI-Systeme sind komplex und intransparent. Das erfordert neue Testansätze. Durch regelmäßige Penetration Tests können Unternehmen ihre Systeme widerstandsfähig machen und somit regulatorische Vorgaben erfüllen. Außerdem stärken sie so das Vertrauen von Partnern und Kunden“, so Hanke weiter.

Generative KI: Innovationskraft mit Verantwortung

TÜV Rheinland unterstützt Unternehmen umfassend beim sicheren Einsatz von KI – von der Durchführung professioneller Penetration Tests über datenbasierte Risikoanalysen bis hin zur Zertifizierung nach international gültigen Standards wie der ISO 42001. „Wer die Chancen generativer KI nutzen will, muss ihrer Sicherheit oberste Priorität einräumen. So lässt sich das Potenzial dieser Technologien verantwortungsvoll erschließen“, betont KI-Experte Daniel Hanke.

Weitere Informationen sowie das Whitepaper unter: www.tuv.com/pentest.

Die Welt zu einem sicheren Ort machen – und das seit mehr als 150 Jahren: Dafür steht TÜV Rheinland als einer der weltweit führenden Prüfdienstleister mit einem Jahresumsatz von mehr als 2,7 Milliarden Euro und 27.000 Mitarbeitenden in gut 50 Ländern. Die hoch qualifizierten Expertinnen und Experten prüfen technische Anlagen und Produkte, begleiten Innovationen und gestalten den Wandel zu mehr Nachhaltigkeit mit. Sie trainieren Menschen in zahlreichen Berufen und zertifizieren Managementsysteme nach internationalen Standards. Mit besonderer Expertise in Mobilität, Energieversorgung, Infrastruktur und vielen weiteren Bereichen sichert TÜV Rheinland unabhängig Qualität, insbesondere bei innovativen Technologien wie grünem Wasserstoff, künstlicher Intelligenz oder automatisiertem Fahren – und ermöglicht so eine sichere und lebenswerte Zukunft. Seit 2006 ist TÜV Rheinland Mitglied im Global Compact der Vereinten Nationen für mehr Nachhaltigkeit und gegen Korruption. Hauptsitz des Unternehmens ist Köln, Deutschland. Website: www.tuv.com

---

Ihr Ansprechpartner für redaktionelle Fragen:
Pressestelle TÜV Rheinland, Tel.: +49 2 21/8 06-21 48
Die aktuellen Presseinformationen sowie themenbezogene Fotos und Videos erhalten Sie auch per E-Mail über contact@press.tuv.com sowie im Internet: www.tuv.com/presse.