AQ SECURE

EMPFEHLUNGEN ZUR ABSICHERUNG KÜNSTLICHER INTELLIGENZ (KI)- UND KI-GESTÜTZTER LÖSUNGEN

AQ SECURE

EMPFEHLUNGEN ZUR ABSICHERUNG KÜNSTLICHER INTELLIGENZ (KI)- UND KI-GESTÜTZTER LÖSUNGEN

1. EINLEITUNG

Künstliche Intelligenz (KI) spielt eine bedeutende Rolle bei der Förderung von Innovation und Effizienz in den Unternehmen und ist inzwischen ein fester Bestandteil datengetriebener Geschäftsprozesse zur Automatisierung, Entscheidungsunterstützung und Optimierung betrieblicher Abläufe.

Unternehmen nutzen zunehmend KI- und KI-gestützte Lösungen zur Auswertung großer und komplexerer Datenmengen, z. B. aus ERP-Systemen, Cloud-Diensten oder SQL-Datenbanken. Diese KI-Lösungen liefern wertvolle Einblicke in Prozesse, Compliance-Risiken oder betriebliche Schwachstellen. Dabei werden oft hochsensible Informationen verarbeitet, die umfassend geschützt werden müssen – sowohl während der Analyse als auch bei der Ergebnisverwendung.

Der Einsatz von Künstlicher Intelligenz zur Optimierung von Geschäftsprozessen bietet ein großes Potenzial, birgt jedoch auch Risiken. Dieses Whitepaper beschreibt Sicherheitsanforderungen für die Implementierung von KI- und KI-gestützten Lösungen, gibt praktische Empfehlungen für technische Schutzmaßnahmen und zeigt, wie Unternehmen die KI- und KI-gestützten Lösungen sicher und regelkonform einsetzen können.

AQ SECURE GmbH
info@aqsecure.de
www.aqsecure.de

AQ SECURE

2. SCHUTZBEDARF SENSIBLER DATEN

Typische Datenarten in Analyseprozessen:

- Personenbezogene Informationen (z. B. Mitarbeitende, Kundendaten)
- Finanztransaktionen (Zahlungen, Kreditoren/Debitoren)
- Workflow- und Freigabedaten
- Audit-Trails und Benutzerverhalten

3. ANGRIFFSSZENARIEN AUF KI-MODELLE

KI-Systeme, insbesondere Machine-Learning-Modelle, sind neuen Angriffsformen ausgesetzt, die über klassische IT-Sicherheit hinausgehen.

ÜBERSICHT DER ANGRIFFSTYPEN UND SCHUTZMASSNAHMEN:

Angriffstyp	**Ziel des Angriffs**	**Beispiel**	**Notwendige Schutzmaßnahmen**
Adversarial Attack	Modell gezielt täuschen (bei der Ausführung)	Manipulierte Eingabe erzeugt Fehlklassifikation	Eingabeüberwachung, robuster Inference-Layer
Data Poisoning	Modell durch Trainingsdaten manipulieren	Vergiftete Daten führen zu Fehlverhalten	Validierung, Monitoring, Source-Härtung
Model Theft	Modell durch API-Zugriffe rekonstruieren	Nachbau durch systematische Abfragen	API-Rate-Limiting, Modell-Wasserzeichen

AQ SECURE GmbH
info@aqsecure.de
www.aqsecure.de

AQ SECURE

4. SCHUTZMASSNAHMEN AUF MODELLEBENE

Die unten empfohlenen Maßnahmen sollen vor allem in produktiven KI-Pipelines greifen, in denen Modelle aus Kundendaten lernen oder automatisierte Entscheidungen treffen.

- Validierung und Bereinigung von Trainingsdaten
- Erkennung adversarialer Eingaben während der Inferenzphase
- Monitoring von Modellverhalten (Drift Detection, Anomalien)
- Ratenbegrenzung bei API-Zugriffen auf ML-Modelle
- Einsatz von Modell-Wasserzeichen und Zugriffskontrollen

5. DATENTRANSPARENZ MIT DATA SECURITY POSTURE MANAGEMENT (DSPM)

Datensicherheitslösungen wie das Data Security Posture Management (DSPM) helfen dabei, sensible und vertrauliche Daten im Unternehmen zu erkennen und zu klassifizieren sowie Risiken frühzeitig zu adressieren, z. B. bevor die Schutzmaßnahmen greifen, die Weitergabe von sensiblen Daten an unbefugten Stellen zu blockieren und Datenverlust zu vermeiden.

TYPISCHE DSPM-KOMPONENTEN:

Komponente	**Funktion**	**Beispiel**
DSPM-Scanner	Identifikation sensibler Datenbestände	Scan von Speicherpfaden, Cloud-Buckets
Risikoanalyse-Dashboard	Visualisierung von Schwachstellen	Aufdeckung falsch konfigurierter Speicher
Policy-Empfehlungen	Ableitung von Schutzregeln	Definition sensibler Datenkategorien

AQ SECURE GmbH
info@aqsecure.de
www.aqsecure.de

AQ SECURE

TYPISCHE DATA SECURITY POSTURE MANAGEMENT (DSPM) EINSATZBEREICHE:

- Verteilte Speicherlandschaften (Hybrid, Multi-Cloud)
- Analyseumgebungen mit unklaren Datenflüssen
- Vorbereitung auf Data Loss Prevention (DLP)-Regelwerke

6. SCHUTZ SENSIBLER DATENBEWEGUNGEN MIT DATA LOSS PREVENTION (DLP)

Um sensible Daten im Unternehmen vor unbefugtem Zugriff, Diebstahl oder Verlust zu schützen, Datenschutzverletzungen zu verhindern, Compliance-Vorschriften einzuhalten sowie finanzielle Reputationsschäden zu minimieren, sollten die Unternehmen zusätzlich zur Data Security Posture Management (DSPM)-Lösung noch eine Implementierung der Data Loss Prevention (DLP)-Lösung berücksichtigen.
Der Einsatz einer Data Loss Prevention (DLP)-Lösung bietet einen umfassenden Ansatz zum Schutz sensibler Daten, der auf der Kombination aus technischen Lösungen, Richtlinien und Mitarbeitersensibilisierung basiert.

TYPISCHE DATA LOSS PREVENTION (DLP)-MODULE:

Modul	**Funktion**	**Beispiel**
Core DLP Engine	Regelbasierte Klassifikation & Kontrolle	Blockieren vertraulicher Dateien
Endpoint DLP Agent	Schutz auf Endgeräten (USB, Clipboard usw.)	Screenshot-Sperre, USB-Kontrolle
E-Mail/Web/Cloud DLP	Kontrolle von Kommunikations- und Upload-Kanälen	Warnung bei sensiblen E-Mail-Anhängen
Risk-Adaptive Protection	Dynamische Reaktion auf Nutzerverhalten	Eskalation bei Anomalien

AQ SECURE GmbH
info@aqsecure.de
www.aqsecure.de

AQ SECURE

TYPISCHE RISIKEN FÜR UNTERNEHMEN OHNE EINSATZ DER DATA LOSS PREVENTION (DLP)-LÖSUNG:

- Weitergabe von Analyseergebnissen an unbefugte Stellen
- Datenabfluss über USB, E-Mail oder Schatten-IT
- Missbrauch durch überprivilegierte Benutzer

7. SCHUTZ DER GENERATIVEN KI- UND KI-GESTÜTZTEN LÖSUNGEN DURCH BROWSER-SICHERHEIT

Da Mitarbeiter typischerweise über den Webbrowser auf generative KI- und KI-gestützte Software-as-a-Service (SaaS)-Lösungen zugreifen, setzen sie Unternehmen auch dem Risiko von Datenverlusten aus, indem sie vertrauliche Unternehmensdaten an die KI-Lösungen weitergeben können. Browser-Sicherheitslösungen können den Unternehmen dabei helfen, integrierte Sicherheitsregeln zu nutzen sowie eigene Richtlinien zu erstellen, um Sicherheits-Governance durchzusetzen, Benutzeraktivitäten bei generativen KI-Lösungen einzuschränken, Datenlecks zu verhindern oder nicht autorisierte KI- und KI-gestützte SaaS-Lösungen zu blockieren.

Mitarbeiter nutzen generative KI-Lösungen (z. B. Text- oder Code-Generatoren) typischerweise über den Webbrowser. Dabei besteht das Risiko, dass sensible Inhalte wie Kundendaten, Quelltexte oder Finanzinformationen unkontrolliert an externe Plattformen weitergegeben werden.

Moderne Browser-Sicherheitslösungen ermöglichen es Unternehmen, diese Risiken gezielt zu steuern:

- Erkennung und Blockierung sensitiver Daten direkt im Browser
- Kontextabhängige Richtlinien zur Nutzung generativer KI (z. B. Schreibschutz, Clipboard-Blocker)
- Protokollierung der Nutzung von Web-KI-Diensten
- Integration mit bestehenden Data Loss Prevention (DLP)- oder Cloud Access Security Broker (CASB)-Systemen

AQ SECURE GmbH
info@aqsecure.de
www.aqsecure.de

AQ SECURE

Solche Funktionen ermöglichen die sichere Nutzung generativer KI – auch bei Bring Your Own AI (BYOAI)-Szenarien oder in offenen SaaS-Umgebungen.

8. SICHERHEITSKULTUR UND AWARENESS

Auch bei technischer Absicherung ist eine starke IT-Sicherheitskultur entscheidend. Schulungen, Sensibilisierung und klare Verantwortlichkeiten helfen, Risiken wie versehentliche Datenweitergabe, Fehlkonfigurationen oder Insider-Missbrauch zu minimieren. Regelmäßige Simulationen, Zugriffsaudits und Transparenz schaffen ein Bewusstsein für sicherheitsrelevante Aspekte bei der Implementierung und Umgang mit KI-Lösungen.

9. UMSETZUNG UND GOVERNANCE

Ein erfolgreicher Schutzansatz für KI- und KI-gestützte Lösungen erfordert:

- Eine schrittweise Einführungsstrategie (Pilot, Skalierung, Betrieb)
- Definierte Rollen (z. B. Datenverantwortliche, Sicherheitsbeauftragte)
- Governance-Vorgaben zur Einhaltung regulatorischer Anforderungen (z. B. DSGVO, ISO 27001, ISO 42001, BSI-Grundschutz)

Eine enge Verzahnung mit bestehenden Sicherheits- und Compliance-Strukturen gewährleistet nachhaltige Sicherheit.

AQ SECURE GmbH
info@aqsecure.de
www.aqsecure.de

AQ SECURE

10. FAZIT

Die erfolgreiche Umsetzung von KI-Initiativen erfordert die Implementierung eines mehrschichtigen KI-Sicherheitskonzepts, das sowohl die Art der KI-Lösung als auch deren Einsatz berücksichtigt und folgende Sicherheitsmaßnahmen enthalten sollte:

- Schutz der KI-Modelle und -Lösungen gegen Manipulation und Extraktion
- Identifikation, Klassifizierung und Risikobewertung sensibler Daten
- Kontrolle der Nutzung und Weitergabe sensibler Daten
- Schutz des Browsers zur Überwachung der Nutzeraktivitäten bei der Nutzung webbasierter generativer KI-Lösungen und zur Blockierung der Weitergabe sensibler Unternehmensdaten im Browser

Durch gezielten Einsatz der KI-, Daten- und Browsersicherheitslösungen lässt sich das volle Potenzial von KI sicher und produktiv nutzen – bei gleichzeitiger Einhaltung von Compliance-Vorgaben und Schutz vor Missbrauch.

Dieses Whitepaper bietet einen technischen Orientierungsrahmen für Unternehmen, die KI-Initiativen sicher und regelkonform gestalten möchten – mit dem Ziel, Innovationspotenziale zu nutzen, ohne Kompromisse bei Datenschutz, Integrität und Governance einzugehen.

Wenn Sie erfahren möchten, wie wir Ihnen helfen können, KI- und KI-gestützte Software-as-a-Service-Lösungen in Ihrem Unternehmen sicher zu entwickeln, einzuführen und zu nutzen, vereinbaren Sie noch heute ein Beratungsgespräch mit unseren Cybersicherheitsexperten.

Wir unterstützen Sie gerne dabei, Ihre KI-Initiativen erfolgreich und sicher zu gestalten.

AQ SECURE GmbH
info@aqsecure.de
www.aqsecure.de