RMPRIVACY

09. April 2024

Richtlinie für den Umgang mit KI-Anwendungen

[Firma]

RMPRIVACY GmbH
Große Langgasse 1a
55116 Mainz

Version 1.3
Freigabe durch: Klicken oder tippen Sie hier, um Text einzugeben. am: Klicken oder tippen Sie hier, um Text einzugeben.
Ablageort: Klicken oder tippen Sie hier, um Text einzugeben.

1 Einleitung

1.1 Diese Richtlinie regelt den Einsatz und die Nutzung von Künstlicher Intelligenz bei der [Firma].

1.2 Dies betrifft insbesondere, aber nicht ausschließlich die Nutzung der Anwendungen ChatGPT, Microsoft Copilot, Leonardo.ai, HeyGen und Dienste, die diese oder andere KI-Anwendungen nutzen.

1.3 Diese Richtlinie soll einen ethischen und verantwortungsvollen Umgang im Einklang mit unseren Unternehmenszielen, aber auch im Umgang mit personenbezogenen Daten im Rahmen von Anwendungen der Künstlichen Intelligenz (KI) sicherstellen.

1.4 Der/die Datenschutzbeauftragte wurde bei der Erstellung dieser Richtlinie beteiligt, und ist bei wesentlichen Änderung zu konsultieren.

1.5 Die Geltung weiterer Richtlinien im Unternehmen bleibt von dieser Richtlinie unberührt.

2 KI und LLMs

2.1 KI bezieht sich auf Technologien und Systeme, die in der Lage sind, Aufgaben auszuführen, die normalerweise menschliche Intelligenz erfordern. Dies umfasst maschinelles Lernen, neuronale Netzwerke, natürliche Sprachverarbeitung und andere verwandte Technologien.

2.2 Ein LLM ist eine bestimmte Art Algorithmus der künstlichen Intelligenz (KI), die maschinelle Lerntechniken und große Datensätze verwendet, um Inhalte zu interpretieren, zusammenzufassen und zu generieren.

3 Ethik und Verantwortung

3.1 Unsere Verwendung von KI-Technologien muss den höchsten ethischen Standards entsprechen. Wir verpflichten uns dazu, sicherzustellen, dass KI-Anwendungen nichtdiskriminierend, fair oder unschädlich sind. Dies schließt die Vermeidung von Vorurteilen und die Gewährleistung von Transparenz ein.

3.2 Wie verpflichten uns, Inhalte kritisch zu prüfen und sicherzustellen, dass sie ethische Standards erfüllen. Die Verbreitung von Fehlinformationen oder schädlichen Inhalten ist zu vermeiden. Wir orientieren uns hier an Leitlinien der Europäischen Kommission. (https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai)

4 Verantwortlichkeiten

Im Rahmen jeder Nutzung von KI-Systemen ist dafür zu sorgen, dass diese sicher und gemäß dieser Richtlinie verwendet werden. Ethische Bedenken sind der Geschäftsführung oder der zuständigen Teamleitung zu melden.

5 IT-Sicherheit & Datenschutz

Es ist sicherzustellen, dass personenbezogene Daten gemäß den geltenden Datenschutzbestimmungen und -richtlinien geschützt werden.

6 Nutzung von personenbezogenen Daten

6.1 Bevor KI-Systeme auf personenbezogene Daten (z.B. Kunden - oder Beschäftigtendaten) angewendet werden, ist eine Risikobewertung, d.h. eine Datenschutz-Folgenabschätzung (DSFA) durchzuführen, um potenzielle Risiken für die Privatsphäre der Kunden zu bewerten.

6.2 Betroffene Personen müssen über die Verwendung von KI in Bezug auf ihre Daten informiert werden und der Verwendung explizit zustimmen, wenn dies erforderlich ist.

6.3 Betroffene haben das Recht, Informationen darüber zu erhalten, wie ihre Daten verwendet werden, und die Möglichkeit, ihre Einwilligung jederzeit zu widerrufen.

6.4 Auf die Eingabe und Nutzung von personenbezogenen und auch anderen sensiblen beziehungsweise vertraulichen Daten sollte bei der Nutzung von ChatGPT und anderen KI-Diensten generell verzichtet werden. Dies gilt auch für Daten Dritter, die in anderen Zusammenhängen verarbeitet wurden.

6.5 Zu vermeiden sind auch Aufforderungen, die zu einer Ausgabe / Widergabe personenbezogener Daten führen können.

7 Prüfung der Ergebnisse

7.1 Bevor Ergebnisse, die mit einer KI erstellt wurden, veröffentlicht oder an einen Kunden gesendet werden, sind diese immer nochmals zu überprüfen und sofern erforderlich und möglich, anzupassen.

7.2 KI-Systeme können irrtümlich bestehende Vorurteile und diskriminierende Muster aus Trainingsdaten übernommen haben. Die Verwertung darauf basierender Ergebnisse würde gegen datenschutzgrundrechtliche Grundsätze, die Vorschriften des Allgemeinen Gleichbehandlungsgesetzes (AGG) und unsere firmeninternen Wertevorstellungen bei [Firma] verstoßen. Bei der Prüfung der Ergebnisse ist hierauf besonders zu achten.

8 Nutzung von KI-Systemen beim Kunden

Bevor KI-Systeme im Zusammenhang mit der Durchführung von Kundenaufträgen genutzt werden, ist dies mit dem Kunden abzustimmen. Dessen Freigabe ist zu dokumentieren.

9 Nutzung dienstlicher E-Mailadressen und Accounts bei Nutzung von KI

KI-Systeme sollen zu dienstlichen Zwecken nur mit den bereitgestellten dienstlichen E-Mailadressen genutzt werden. Sofern Funktionsaccounts ohne Personenbezug bereitgestellt werden, sind diese zu nutzen.

10 Keine automatisierten Letztentscheidungen

Entscheidungen, die gegenüber natürlichen Personen Rechtswirkungen entfalten oder andere Nachteile nach sich ziehen können (etwa: Entscheidung über Einstellungen, Kündigungen, Vertragsabschlüsse) sind von Menschen zu treffen.

11 Technische Konfigurationen bei Anwendung von KI

11.1 Keine Nutzung zu Trainingszwecken

11.1.1 Der Anbieter von ChatGPT behält sich ausdrücklich die Nutzung der eingegeben Daten zu Trainingszwecken der KI vor. Diese Möglichkeit ist innerhalb der Einstellungen (unter dem Reiter „Data controls“) zu deaktivieren.

Eine Hilfestellung zum Abschalten der Trainingsmöglichkeit von ChatGPT findet sich unter: https://help.openai.com/en/articles/5722486-how-your-data-is-used-to-improve-model-performance.

11.1.2 Auch bei anderen Anbietern sind KI-Tools dergestalt zu konfigurieren, dass eine Weitergabe von Daten zu Trainingszwecken nicht zugelassen ist.

11.2 Aufbau einer Zero-Trust-Architektur

11.2.1 Zero Trust ist eine Strategie, die die Sicherheit von Systemen gewährleisten soll. Sie beruht auf den folgenden drei Prinzipien:

(a) Explizit verifizieren: Ziehen Sie zur Autorisierung und Authentifizierung alle verfügbaren Datenpunkte heran. Anwendungen oder Anwender sind nicht automatisch deswegen vertrauenswürdig, weil sie aus dem eigenen Netzwerk stammen („hinter der FireWall“), sondern Berechtigungen sind immer zu prüfen.

(b) Verwenden des Zugriffs mit den geringsten Rechten: Benutzern ist nur der benötigte Zugriff und die benötigte Zeit um ihre Aufgaben auszuführen, im System zu gewähren.

(c) Von einer Sicherheitsverletzung ausgehen, und den Standard der Maßnahmen hieran orientieren.

11.2.2 Bevor KI-Dienste, insbesondere Microsoft Copliot im Unternehmen implementiert werde, ist vorab eine Zero-Trust-Architektur im Unternehmen aufzubauen. Microsoft stellt entsprechende Hilfestellungen zur Konfiguration seiner Produkte hier zur Verfügung: https://learn.microsoft.com/de-de/security/zero-trust/

11.3 Einrichtung von Funktionsaccounts

Sofern möglich, sind für die Nutzung von KI-Diensten Funktionsaccounts zur Verfügung zu stellen, die über das Unternehmen registriert sind. Bei der Erstellung der Accounts sind die Nutzungsbedingungen der Anbieter zu berücksichtigen, insbesondere ob eine Registrierung ohne Angabe personenbezogener Daten zulässig ist.

11.4 Regelmäßige Löschung

Sofern eine Löschung der Eingaben und Verläufe nicht von den Mitarbeitern selbst vorgenommen werden kann oder wird, sind administratorseitig regelmäßige Löschungen der bisherigen Verlaufsdaten und Eingaben zu veranlassen oder einzustellen, spätestens alle 180 Tage.