KI-Richtlinie des Bezirksamtes Mitte von Berlin
(RL-KI-BAMi)

Änderungshistorie

erstellt am	24.04.2024	von	ISB
ergänzt am	18.06.2024	von	ISB
überprüft am	19.06.2024	von	E-Government-Team
ergänzt am	03.07.2024	von	ISB
ergänzt am	29.08.2024	von	ISB
		von	
		von	
		von	
		von	
		von	
		von	

Version: 0.04

Status: zur Freigabe

Inhaltsverzeichnis

1 Einleitung 3
2 Geltungsbereich 3
3 Schutzbedarfsfeststellung – Was ist KI und welche Chancen und Risiken ergeben sich daraus? 3
4 Maßnahmen – Was ist erlaubt? Was ist verboten? 4
5 Wohin mit Fragen? 6
6 Sicherheitsorganisation 6
7 Rollen 6
7.1 Informationssicherheitsbeauftragter 6
7.2 Datenschutzbeauftragte 6
7.3 IKT-Management 6
7.4 E-Government-Team 6
8 Folgen von Zuwiderhandlung 7
9 Inkrafttreten 7

1 Einleitung

Nach der Einführung von ChatGPT durch OpenAI im November 2022 wurde das Thema der frei verfügbaren, hochentwickelten künstlichen Intelligenzen deutlich populärer. Plötzlich kann jeder und jede mit sehr geringem Aufwand und Vorwissen mit einer künstlichen Intelligenz interagieren. Regelungen in der öffentlichen Verwaltung fehlen bislang. Umso wichtiger ist es daher, dass intern Regeln für die Nutzung von künstlichen Intelligenzen bereitstehen. Diese Richtlinie klärt darüber auf, was künstliche Intelligenz ist, welche Chancen und Risiken bestehen und was bei einer Nutzung zu beachten ist. Bedenken Sie, dass die Nutzung freiwillig, ähnlich einer Suchmaschine (Google, Ecosia, etc.) erfolgt. Es ist nicht notwendig oder vorgeschrieben, diese Tools für die tägliche Arbeit zu verwenden.

2 Geltungsbereich

Diese Richtlinie gilt uneingeschränkt für die dienstliche Nutzung von KI für Tätigkeiten, die die Beschäftigten des Bezirksamtes Mitte von Berlin im Rahmen ihres Arbeitsverhältnisses ausüben. Sie muss standortunabhängig eingehalten werden. Insbesondere die Maßnahmen in Punkt 4 müssen durchgehend gewährleistet sein.

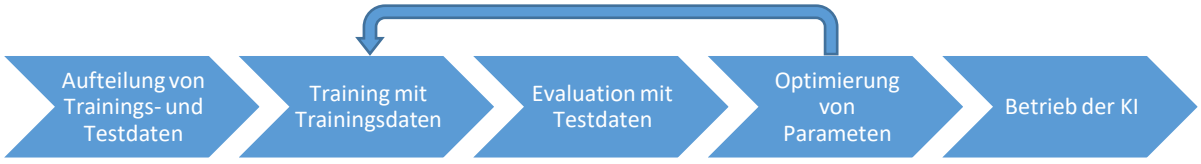
Mitgeltende Dokumente:

- Informationssicherheitsleitlinie des Landes Berlin (Version 1.0.1)
- IKT-Architektur (in der jeweils gültigen Version)
- Handout der SenFin zur Nutzung von KI vom 21.02.2024 (Version 1.0)
- Informationssicherheitsleitlinie des Bezirksamtes Mitte von Berlin (Version 1.4)
- Datenschutzkonzept des Bezirksamtes Mitte von Berlin (Version 1.0.8 F)

3 Schutzbedarfsfeststellung – Was ist KI und welche Chancen und Risiken ergeben sich daraus?

KI (Künstliche Intelligenz) beschreibt mathematische Modelle, die trainiert wurden, um Aufgaben des alltäglichen Lebens zu vereinfachen. So können diese Modelle zum Beispiel Texte erschaffen oder zusammenfassen, Bilder und Videos auf Grundlage einer Beschreibung erzeugen, Sprache verstehen, übersetzen oder ganze Stimmen imitieren. KI hat dennoch kein Bewusstsein.

Um eine KI zu erstellen wird vereinfacht das folgende Schema durchlaufen:



Eine KI ist dabei immer nur so gut, wie ihr Trainingsdatensatz war. Wenn der Trainingsdatensatz Falschinformationen enthielt, dann kann die KI Falschinformationen ausgeben (wenn das gezielt passiert nennt sich das Data Poisoning: Vergiften von Daten). Es ist einer KI auch möglich, sich „Fakten“ auszudenken. ChatGPT wurde über alle frei verfügbaren Inhalte des Internets trainiert (Also bspw. jede Seite, die Sie bei Google finden können wurde einmal geöffnet, analysiert und gespeichert). Dabei sind auch kreative Texte eingelesen worden, welche keinen Bezug zur Realität haben. Manche KI-Dienste haben keinen Zugriff auf Echtzeitdaten. Sie wurden also mit bestehenden Daten trainiert, erhalten aber keine neuen Informationen.

An dieser Stelle ergeben sich drei Risiken:

1. Die KI kann Dinge als Fakten darstellen, die vollständig ausgedacht sind.
2. Außerdem können dargestellte Informationen veraltet sein
3. Gleichzeitig kann das Copyright von Künstlerinnen und Künstlern verletzt werden, indem Teile von Werken als Arbeit der KI ausgegeben werden.

Beachten Sie an dieser Stelle auch, dass KI unethisch handeln kann. Ethik beschäftigt sich vereinfacht gesagt mit der Grundsatzfrage: Was ist gut/richtig und was ist schlecht/falsch? KI-Anwendungen sind nur so gut, wie die Trainingsdaten. Wenn diese Verzerrungen enthalten, weil die Personen, welche die Entscheidungen getroffen haben, nicht vorurteilsfrei waren, dann wird das trainierte System diese Verzerrungen/Diskriminierungen, etc. auch aufweisen.

Die Ausgabe von KI-Systemen ist daher immer zu hinterfragen und nicht blind anzunehmen.

Während der Trainingsphase benötigt das mathematische Modell extrem viel Rechenleistung. Um diese Berechnungen in annehmbarer Zeit durchzuführen werden Hochleistungsserver verwendet.

Leider sind diese auch nach Abschluss des Trainings noch notwendig um in kurzer Zeit eine umfangreiche Rückmeldung der KI zu erhalten. Derzeit sind nahezu alle frei verfügbaren künstlichen Intelligenzen ausschließlich auf Servern außerhalb des ITDZ und des Bezirksamtes Mitte verortet. Wir können also nicht beeinflussen, welche (personenbezogenen) Daten auf diesen Servern verarbeitet (gesammelt, gespeichert, weitergegeben) werden.

4 Maßnahmen – Was ist erlaubt? Was ist verboten?

Grundsatz: Alles was frei im Internet verfügbar ist, darf auch als Eingabe für KI-Systeme genutzt werden. Alles andere ist verboten.

✓ Einsatz von KI erlaubt	✗ Einsatz von KI untersagt
Zusammenfassung längerer, öffentlich-zugänglicher Texte (Copyright beachten!)	geheime, interne oder unveröffentlichte Texte zusammenfassen lassen
Definitionen abfragen und um einfache Beispiele aus der realen Welt ergänzen	Ergebnisse der KI ungeprüft übernehmen (Kontrolle: Ergebnisse stets auf Richtigkeit und Vollständigkeit, potentielle Diskriminierung

	überprüfen, Frage durch unterschiedliche Eingabe validieren)
Bilder und Visualisierungen ggf. auch die Struktur für Präsentationen erstellen lassen (Eingabe darf nur öffentlich zugängliche Informationen beinhalten)	erzeugte Bilder ohne vorherige Prüfung auf potentielle Urheberrechtsverletzung verwenden (Verwendung von Markennamen)
Inhalte von https://daten.berlin.de und https://berlin.de	Inhalte aus dem Intranet (b-intern) dürfen nicht für Eingaben der KI genutzt werden.

Kurze Auffrischung zu personenbezogenen Daten: Darunter werden alle Daten verstanden, welche dazu beitragen, eine Person eindeutig identifizieren zu können. Beispiele sind: Name, Anschrift, Telefonnummer, E-Mail-Adresse, Webseite, IP-Adresse, Fingerabdrücke, Informationen über den Gesundheitsstatus, Glaubensstatus, etc. Personenbezogene Daten erhalten Sie ausschließlich über interne Informationen. **Abfragen an die KI, die personenbezogene Daten beinhalten sind somit untersagt.**

Dazu zählen auch unstrukturierte Daten. Folgendes wäre ebenfalls eine Eingabe mit Personenbezug: „Entwirf ein Arbeitszeugnis im befriedigenden Bereich für einen Kundenberater im Autohaus X."

Sie wollen KI für dienstliche Zwecke nutzen, wie starten Sie?

- Legen Sie ein Konto bei der entsprechenden KI-Dienstseite an
- Nutzen Sie dazu Ihre dienstliche E-Mail-Adresse. Bitte seien Sie sich bewusst, dass der KI-Dienst Ihre E-Mail-Adresse und somit auch Ihre personenbezogenen Daten (mindestens E-Mail, Vorname, Nachname) erhält. Wenn Sie das nicht wollen, steht es Ihnen frei den Dienst nicht zu verwenden oder Sie nutzen eine anmeldefreie Variante von DuckDuckGo: https://duckduckgo.com/?q=DuckDuckGo&ia=chat.
- **Beachten Sie, dass die IT-Stelle Ihnen bei Fragen zum Passwort oder Benutzereingaben keine Unterstützung geben kann.** Auch Anleitungen stehen nicht zentral zur Verfügung.
- Verwenden Sie ein langes Passwort (erzeugen Sie dies bspw. mit dem Passwortmanager KeePassXC, den Sie sich von der IT-Stelle zuweisen lassen können)
- Offene Daten mit Bezug zur Verwaltung können Sie bspw. auf https://daten.berlin.de oder auf den Webseiten von https://berlin.de einsehen.
- **Wenn Sie sich unsicher sind: Die Nutzung ist auf freiwilliger Basis. Sie müssen die KI-Werkzeuge nicht verwenden.**

5 Wohin mit Fragen?

Bei Fragen zum Datenschutz und zur Informationssicherheit können Sie sich an folgende Personen wenden:

Sandra Müller
Datenschutzbeauftragte
datenschutz@ba-mitte.berlin.de

Jakob Küchler
Informationssicherheitsbeauftragter
it-sicherheit@ba-mitte.berlin.de

6 Sicherheitsorganisation

Diese Richtlinie muss im Rahmen der regelmäßigen Überprüfung des ISMS ebenfalls geprüft und gegebenenfalls aktualisiert werden. Die Federführung unterliegt dabei dem/der Informationssicherheitsbeauftragten in Abstimmung mit dem E-Government-Team.

7 Rollen

7.1 Informationssicherheitsbeauftragter

Der Informationssicherheitsbeauftragte ist für die Aktualisierung und Fortschreibung dieser Richtlinie in Zusammenarbeit mit dem E-Government-Team verantwortlich.

7.2 Datenschutzbeauftragte

Die Datenschutzbeauftragte setzt sich für den Schutz aller personenbezogenen Daten ein. Sie beantwortet alle Fragen zu diesem Fachgebiet und wirkt im Rahmen des E-Government-Teams an der Fortschreibung dieser Richtlinie mit.

7.3 IKT-Management

Das IKT-Management bildet die Leitungsebene von IT-Stelle, TK und technischer Beschaffung und ist als Mitglied des E-Government-Teams an der Fortschreibung und Aktualisierung dieser Richtlinie beteiligt.

7.4 E-Government-Team

Das E-Government-Team besteht gemäß BA Beschluss 489 vom 24.07.2018 regelmäßig aus Informationssicherheitsbeauftragtem, IKT-Management, Business Continuity Management und Steuerungsdienst. Anlassbezogen werden die Datenschutzbeauftragte, Rechtsamt und Pressestelle hinzugezogen. Das E-Government-Team arbeitet bei der Fortschreibung und Aktualisierung dieser Richtlinie zu. Dazu können ggf. weitere OE hinzugezogen werden.

8 Folgen von Zuwiderhandlung

Vorsätzliche oder grob fahrlässige Handlungen, die Sicherheitsvorgaben verletzen, können finanzielle Verluste bedeuten, Beschäftigte, Geschäftspartner*innen und Kunden schädigen oder den Ruf des Bezirksamtes gefährden. Bewusste Verstöße gegen verpflichtende Sicherheitsregeln können arbeitsrechtliche, disziplinarrechtliche und unter Umständen auch strafrechtliche Konsequenzen haben. Zivilrechtliche Regressforderungen sind möglich.

9 Inkrafttreten

Diese Richtlinie tritt zum XX.XX.2024 in Kraft

Freigegeben durch: BA-Beschluss XX.