Merkblatt
Datenschutz und Nutzung generativer KI-Anwendungen durch öffentliche Organe

Allgemeines

Die folgenden Informationen beziehen sich auf die datenschutzrechtlichen Vorgaben im Zusammenhang mit dem **Einsatz von internetbasierten Anwendungen mit generativer künstlicher Intelligenz (KI)**, beispielsweise ChatGPT von OpenAI, Copilot von Microsoft, Gemini von Google, Grok von X etc. im Arbeitskontext der öffentlichen Verwaltung.

Die Bearbeitung von **Personendaten** oder anderen **sensitiven Informationen** mit entsprechenden Diensten ist **nicht erlaubt**.

Zulässige Nutzung

Die Verwendung generativer KI ist aus Datenschutzsicht **beispielsweise für folgende Zwecke unproblematisch:**

- Zusammenfassung und sprachliche Überarbeitung öffentlich zugänglicher Informationen,
- Unterstützung bei Strukturierung und Formulierung von Texten, sofern keine Personendaten oder andere sensitive Informationen involviert sind,
- allgemeine Recherche und Einarbeitung in ein Thema.

Die Verwendung ist jedoch auch dann mit **Risiken** verbunden. Insbesondere:

- Die Entstehung von Ausgabeergebnissen ist nicht nachvollziehbar und intransparent
- Die KI generiert unzutreffende oder erfundene Ergebnisse
- Texteingaben können durch die Anbieter der KI zu eigenen Trainingszwecken verwendet werden
- Unbeabsichtigte Eingabe von Personendaten oder anderer sensitiver Informationen
- Aufgrund schlechter Trainingsdaten werden verzerrte und diskriminierende Ergebnisse erzeugt (Biases)

Unzulässige Eingaben und Verwendung

Die Eingabe von Personendaten ist grundsätzlich untersagt. Dies gilt insbesondere für:

- Angaben zu identifizierten oder identifizierbaren natürlichen Personen (z.B. Namen von Bürgerinnen oder Mitarbeitenden, E-Mail-Adressen oder Telefonnummern, Fallbeschreibungen mit Bezug zu einer Person),
- besonders schützenswerte Personendaten (z.B. Gesundheits- oder Sozialhilfedaten, Angaben zu Straftaten und Strafen und Massnahmen)
- Daten über administrative, disziplinarische oder rechtliche Verfahren,
- Kombinationen von Einzelangaben, die eine Identifikation ermöglichen.

Auch **pseudonymisierte oder anonymisierte Daten** dürfen nicht eingegeben werden, sofern eine Re-Identifikation nicht praktisch vollständig ausgeschlossen werden kann. Das einfache Entfernen oder Ersetzen identifizierender Merkmale genügt nicht, wenn durch Kombination mit anderen (frei zugänglichen) Informationen die Person bestimmbar wird.

Ebenfalls unzulässig ist die Eingabe von anderen sensitiven Informationen. Dazu zählen insbesondere:

- interne oder vertrauliche Informationen,
- Daten, die dem Amtsgeheimnis oder besonderen (Berufs-) Geheimnispflichten (z.B. Sozialhilfegeheimnis, Steuergeheimnis, Arztgeheimnis) unterliegen,
- vertraglich geschützten Informationen,
- sicherheitsrelevanten System- oder Infrastrukturangaben.

Transparenz und automatisierte Entscheidungen

Entscheidungen dürfen nicht ausschliesslich automatisiert erfolgen, wenn sie für betroffene Personen eine Rechtsfolge haben oder sie erheblich beeinträchtigen, sofern die gesetzlichen Voraussetzungen nicht erfüllt sind.

Betroffene Personen haben Anspruch auf **Information** über eine automatisierte Bearbeitung. Die Nutzung von KI ist gegebenenfalls transparent zu machen.

Prüf- und Sorgfaltspflicht

Generative KI-Systeme erstellen Inhalte auf Basis statistischer Wahrscheinlichkeitsmodelle. Die erzeugten Resultate können sachlich falsch, unvollständig, veraltet oder verzerrt sein. Alle durch KI erzeugten Inhalte sind vor Verwendung sorgfältig auf Richtigkeit, Vollständigkeit, Aktualität und rechtliche Zulässigkeit zu **prüfen**.

Entscheide müssen jederzeit sachlich nachvollziehbar und eigenständig begründet werden können. Werden Ergebnisse generativer KI in Entscheidungsprozesse einbezogen, bleibt die **Verantwortung** vollumfänglich bei der zuständigen Mitarbeiterin bzw. beim zuständigen Mitarbeiter.

Weiteres

Bezüglich **Anschaffung oder Einsatz von KI-Tools in der Verwaltung**, wie auch bei der Nutzung oder Anschaffung von nicht-KI-basierter Technologie, sind vorab die **datenschutzrechtlichen Vorgaben zu konsultieren** und deren **Einhaltung sicherzustellen**. Die Nutzung eines KI-Systems darf erst dann erfolgen, wenn die **datenschutzrechtlichen Voraussetzungen** erfüllt und vorab allfällige **Risiken** angemessen beurteilt und adressiert wurden. Sollen mit einer KI-Anwendung Personendaten bearbeitet werden, ist eine **Datenschutz-Folgenabschätzung (DSFA)** und allenfalls eine **Vorabkonsultation** durchzuführen. Wird der Einsatz einer KI-Anwendung geplant, die mit besonders risikobehafteten Datenbearbeitungen verbunden ist (z.B. Einsatz im Bereich Steuer-Veranlagung), ist eine Regelung in einem **Gesetz** vorzusehen.

Auch die geltenden Vorgaben zur Informationssicherheit sowie zur Informatik- und Cybersicherheit sind jederzeit einzuhalten.