Gemeindeverwaltung Allschwil

Verwaltungsführung
Rechtsdienst

Weisung zur Nutzung von generativen KI-Werkzeugen

1 Was sind generative KI-Werkzeuge?

Im Internet verfügbare Werkzeuge mit generativer künstlicher Intelligenz (KI) – zum Beispiel ChatGPT von OpenAI, Copilot von Microsoft, Bard von Google, Grok von X und zahlreiche mehr – vereinfachen eine Reihe von Aufgaben, die auch in der Verwaltung zum Arbeitsinhalt vieler Mitarbeitenden gehören. Sie ermöglichen es den Nutzenden, beispielsweise die KI-Werkzeuge um eine Stellungnahme zu einem bestehenden Text zu bitten oder sie aufzufordern, einen neuen Text zu einem bestimmten Thema zu erstellen. Diese Werkzeuge sind nicht «intelligent»; sie berechnen z.B. bei der Textgenerierung lediglich die statistische Wahrscheinlichkeit der Wortteilfolge – sie sind also next token prediction systems – liefern aber dennoch oft erstaunliche Ergebnisse. Sie werden mit grossen Datenmengen gefüttert, deren Quellen meistens nicht offengelegt sind. Die darauf berechneten Wahrscheinlichkeiten können daher veraltet, irreführend, diskriminierend oder schlicht falsch sein. Ebenso dienen die Eingaben (sog. Prompts oder Eingabeaufforderungen) unter Umständen dem weiteren Training des KI-Systems, sie können also in andere Unterhaltungen einfliessen. Die Daten können je nach KI-Werkzeug auch ausserhalb der Schweiz gespeichert werden.

2 Ziel der Weisung

Diese Weisung hat das Ziel, den Missbrauch von externen KI-Werkzeugen zu verhindern und sicherzustellen, dass deren Nutzung im Einklang mit ethischen, datenschutzrechtlichen und gesetzlichen Vorgaben erfolgt. Dabei werden der Schutz der Rechte betroffener Personen, die Sicherstellung von Transparenz sowie die Vermeidung von Diskriminierung und ungerechtfertigter Benachteiligung berücksichtigt.

3 Geltungsbereich

Diese Weisung gilt für alle Mitarbeitenden, die externe oder intern zur Verfügung gestellte KI-Werkzeuge im Rahmen ihrer beruflichen Tätigkeit nutzen. Sie bezieht sich auf die Nutzung dieser Werkzeuge in allen Abteilungen, die in irgendeiner Weise Entscheidungen treffen, die durch KI-Systeme beeinflusst werden oder Personendaten auf diese Weise verarbeiten.

4 Begriffe (gem. Art. 5 DSG, § 3 IDG BL)

4.1 Personendaten

Personendaten umfassen alle Informationen, die sich auf eine bestimmte oder bestimmbare natürliche Person beziehen.

4.2 Anonymisieren

Anonymisierte Daten sind Angaben, bei denen der ursprünglich vorhandene Personenbezug dauerhaft entfernt worden ist. Sie sind keine Personendaten mehr, wenn die Re-Identifikation nicht oder nur mit einem unverhältnismässigen Aufwand möglich ist

4.3 Bearbeiten

Jeder Umgang mit Informationen, unabhängig von den angewandten Mitteln und Verfahren, insbesondere das Beschaffen, Aufbewahren, Lesen, Verändern, Bekanntgeben, Archivieren, Löschen oder Vernichten sowie das Durchführen logischer und/oder rechnerischer Operationen mit diesen Informationen.

5 Grundsätze

5.1 Zweckbindung

Jede Erhebung von Personendaten setzt den Zusammenhang mit der Erfüllung einer öffentlichen Aufgabe voraus. Personendaten dürfen nur zu dem Zweck bearbeitet werden, zu dem sie erhoben worden sind, soweit nicht eine gesetzliche Grundlage ausdrücklich eine weitere Verwendung vorsieht oder die betroffene Person im Einzelfall einwilligt.

5.2 Transparenz

Treffen Mitarbeitende Entscheidungen, welche auf Ergebnissen von generativen KI-Werkzeugen beruhen, so müssen sie diese jederzeit begründen können. Bei Bedarf muss es möglich sein, die Entscheidung des KI-Systems zu hinterfragen und gegebenenfalls zu korrigieren.

5.3 Verantwortung

Mitarbeitende tragen die Verantwortung für alle Entscheidungen und Ergebnisse, die auf den Ergebnissen des externen KI-Werkzeugs basieren. Mitarbeitende, welche zur Erfüllung ihres gesetzlichen Auftrags Informationen bearbeiten, sind für den korrekten Umgang mit den Informationen verantwortlich. Selbst wenn das KI-Werkzeug Fehler oder Vorurteile aufweist, bleibt die Verantwortung bei den Mitarbeitenden.

6 Datenschutz

Es dürfen keine unbefugten Daten verarbeitet oder weitergegeben werden. Dabei sind insbesondere Personendaten zu schützen. Personendaten sind zu anonymisieren, bevor sie in das externe KI-Werkzeug eingegeben werden.

7 Vorgaben zur Nutzung

7.1 Mögliche Einsatzbereiche

Längere öffentlich verfügbare Texte lassen sich zusammenfassen, die Struktur einer Präsentation kann optimiert werden, Vorschläge für E-Mails, Texte oder Berichte können erstellt werden, oder es ist möglich, sich effizient und interaktiv in ein neues Thema einzuarbeiten.

7.2 Missbrauchsvermeidung und ethische Nutzung

Externe oder intern zur Verfügung gestellte KI-Werkzeuge dürfen nur zu ethisch vertretbaren Zwecken eingesetzt werden. Die Nutzung des Systems zu diskriminierenden, manipulativen oder missbräuchlichen Zwecken ist strikt untersagt. Insbesondere dürfen keine unzulässigen oder vorgefassten Entscheidungen getroffen werden, die einzelne Personen oder Gruppen benachteiligen.

Es dürfen niemals vertrauliche, als intern klassifizierte, persönliche Daten oder sensible Informationen in diese Werkzeuge eingegeben werden.

- keine Eingabe von als intern, vertraulich oder geheim klassifizierten Informationen;
- keine Eingabe von Texten, die zwar nicht klassifiziert sind, aber sensible Informationen enthalten, etwa weil sie durch eine Geheimhaltungspflicht geschützt sind (Amtsgeheimnisse, besondere Berufsgeheimnisse, vertraglich ausdrücklich geschützte Informationen). Vorsicht auch bei der Eingabe von Bildern, privaten Fotos, Tonaufnahmen, Videos, Simulationen und Code;
- keine Eingabe von Personendaten jeglicher Art. Bei der Eingabe von anonymisierten oder pseudonymisierten Daten ist darauf zu achten, dass nicht aufgrund von zusätzlichen Informationen doch Rückschlüsse auf die Betroffenen gezogen werden können (etwa indem zwar ein Name abgeändert wird, aber aufgrund der Angabe des Geburtsdatums, des Geschlechts und des Wohnquartiers die fragliche Person relativ einfach in Erfahrung gebracht werden kann).

Bei Unklarheit ob der Qualifikation der zu verwendenden Informationen und Daten, soll auf deren Eingabe und die Benutzung der generativen KI-Werkzeuge verzichtet werden. Bereits öffentlich (im Internet) publizierte Informationen, wie Open Government Data (OGD), dürfen verwendet werden.

7.3 Regelungen bei Fehlern

Bei der Feststellung von fehlerhaften, diskriminierenden oder unerwünschten Ergebnissen aufgrund der Nutzung eines externen KI-Werkzeugs sind diese zu verwerfen. Mitarbeitende dürfen das KI-Werkzeug nicht in einer Weise verwenden, die darauf abzielt, Fehler zu verschleiern oder verfälschte Daten einzugeben, um Ergebnisse zu beeinflussen.

7.4 Vergleich mit anderen Quellen

Generative KI-Werkzeuge liefern Ergebnisse unterschiedlicher Qualität. Die Ergebnisse der Werkzeuge sind im Zweifelsfall kritisch auf Richtigkeit und Vollständigkeit zu überprüfen und mit anderen Quellen zu vergleichen.

7.5 Private Nutzung

Die von der Gemeinde zur Verfügung gestellten KI-Werkzeuge dürfen ausschliesslich für die Erfüllung der beruflichen Aufgaben genutzt werden. Eine Nutzung für private Zwecke ist untersagt.

7.6 Einfache Suchanfragen

Aus Nachhaltigkeitsgründen und zur Unterstützung übergeordneter Klimaziele sollten für einfache Recherchen bzw. Suchanfragen weiterhin die gängigen Such-Portale (wie z.B. Google oder Microsoft Bing) verwendet werden.

8 Zugriffsrechte

Der Zugriff auf externe KI-Werkzeuge und die darin verarbeiteten Daten können bei Bedarf für die Mitarbeitenden eingeschränkt oder gesperrt werden.

9 Konsequenzen bei Verstössen

Verstösse gegen diese Weisung, insbesondere in Bezug auf Datenschutz, Missbrauch oder unrechtmässige Nutzung des KI-Werkzeugs, werden entsprechend den Vorgaben im Personal- und Besoldungsreglement geahndet. Abhängig von der Schwere des Verstosses können dies Abmahnungen, Kündigungen oder rechtliche Schritte gegen die verantwortliche Person beinhalten.

10 Gültigkeit und Änderungen

Diese Weisung tritt am 1. März 2025 in Kraft. Sie wird regelmässig überprüft und wird bei Bedarf an neue gesetzliche Anforderungen oder technologische Entwicklungen angepasst.

Patrick Dill

Leiter Gemeindeverwaltung

Jesse van Rijswijk

Bereichsleiter Finanzen – Informatik – Personal

Versionenverlauf			
Version	**Anpassungen / Ergänzungen**	**Autoren**	**Datum**
1.0	Initiale Version durch GL genehmigt	Mirjam Glarner / Jesse van Rijswijk	20.02.2025