



Richtlinie der Stadeskommission zur Nutzung von generativen KI-Werkzeugen

vom 1. Oktober 2024

1 Ausgangslage

Generative KI-Werkzeuge sind öffentlich zugängliche Online-Anwendungen, welche basierend auf einer Eingabe mit generativer Künstlicher Intelligenz (KI) automatisch Texte, Bilder, Programmcodes und andere Datenformate als Ausgabe generieren. ChatGPT, Google Bard und Microsoft Copilot sind Beispiele von generativen KI-Werkzeugen.

Die Mitarbeitenden der kantonalen Verwaltung Appenzell I.Rh. können generative KI-Werkzeuge grundsätzlich unter Beachtung der geltenden rechtlichen Vorgaben im Arbeitsalltag verwenden. Im Umgang mit generativen KI-Werkzeugen ist allerdings generell Vorsicht geboten.

2 Richtlinie

Die Richtlinie ist auf die Nutzung von generativen KI-Werkzeugen zur Erstellung von Texten in der Verwaltung und für Verwaltungsmitarbeitende ausgerichtet. Die Grundsätze gelten sinngemäss für alle Arten von generativen KI-Werkzeugen.

2.1 Nutzung von generativen KI-Werkzeugen

Generative KI-Werkzeuge eignen sich vorwiegend als Arbeitshilfe zum persönlichen Gebrauch, beispielsweise

- um eine Zusammenfassung von öffentlich zugänglichen Informationen zu erstellen,
- um Ideen zu einem allgemeinen Thema zu finden,
- als Formulierungshilfe für allgemeine eigene Texte oder für E-Mails ohne Personendaten und vertrauliche Inhalte.

Für viele Verwaltungsaufgaben sind generative KI-Werkzeuge nicht geeignet. Auf den Einsatz von generativen KI-Werkzeugen ist in folgenden Fällen zu verzichten:

- Beim Vorbereiten von rechtsverbindlichen Entscheiden.
- Beim Erarbeiten von Arbeitsdokumenten auf der Basis von verwaltungsinternen Informationen.
- Im Austausch mit der Bevölkerung, beispielsweise bei der Beantwortung von Anfragen, sofern es sich nicht um öffentlich zugängliche Informationen handelt.

2.2 Eingabe von Informationen im Eingabefeld

Werden generative KI-Werkzeuge genutzt, ist bei der Eingabe von Informationen (prompt) auf Folgendes zu achten:

- Die Eingabe darf keine Personendaten enthalten.
- Die Eingabe darf keine als intern oder vertraulich klassifizierten Informationen enthalten.

- Die Eingabe darf keine Information enthalten, die einer Organisationseinheit zugeordnet werden kann. Dazu gehören beispielsweise Amtsbezeichnungen, interne Kennungen oder Aktenzeichen.

2.3 Weiterverwendung der Ergebnisse

Die Ausgabe eines generativen KI-Werkzeugs basiert auf Wahrscheinlichkeitsberechnungen, was bedeutet, dass das Ergebnis nur rechnerisch wahrscheinlich und nicht zwingend inhaltlich korrekt ist. Bei der Weiterverwendung der Ergebnisse von generativen KI-Werkzeugen ist daher Folgendes zu beachten:

- Generierte Texte und die darin enthaltenen Informationen sind kritisch auf Qualität und Korrektheit zu prüfen.
- Um die Richtigkeit eines Ergebnisses sicherzustellen, sollte ein Vergleich mit anderen, zuverlässigen Quellen vorgenommen werden.
- Generierte Texte sind sachlich anzupassen und an die jeweils geltenden Schreibweisungen anzugleichen. Insbesondere sind das [Manual](#) und die einschlägigen [kantonalen Merkblätter](#) zu berücksichtigen.
- Die Ergebnisse von generativen KI-Werkzeugen sind oft nicht neutral oder enthalten Wertungen. Die Texte sind kritisch zu prüfen und zu bereinigen. Insbesondere ist zu prüfen, ob ein Text diskriminierend wirkt.
- Die Ergebnisse von generativen KI-Werkzeugen basieren auf bestehenden Texten. Diese können urheberrechtlich geschützt sein.

2.4 Rechtliche Vorgaben

Bei der Nutzung von generativen KI-Werkzeugen sind die allgemeinen rechtlichen Vorgaben bezüglich Vertraulichkeit, Informationssicherheit und Datenschutz zu beachten.

Zu berücksichtigen sind insbesondere:

- der Schutz von Personendaten (siehe Datenschutz-, Informations- und Archivgesetz, DIAG, GS 172.800, und Art. 13 Bundesverfassung, SR 101),
- die Sicherheit von Daten,
- die Wahrung des Amtsgeheimnisses (siehe Schweizerisches Strafgesetzbuch, SR 311.0),
- personalrechtliche Pflichten (siehe Personalverordnung, PeV, GS 172.310),
- spezifische Vorgaben über die Nutzung von Informatikmitteln (siehe Ständekommissionsbeschluss über die Informatiknutzung, GS 172.315),
- bestehende Urheberrechte (Urheberrechtsgesetz, SR 231.1).

2.5 Empfehlung

Beim Einsatz von generativen KI-Werkzeugen wird die Verwendung von «Microsoft Copilot mit kommerziellem Datenschutz» empfohlen. Bei dessen Nutzung werden die eingegebenen Informationen nur temporär gespeichert und nicht für das Trainieren des Modells verwendet. Zudem wird bei den Ergebnissen auf die zugrundeliegenden Quellen verwiesen. Eine im Intranet angeschaltete [Anleitung](#) erläutert den Anmeldevorgang für die Nutzung von «Microsoft Copilot mit kommerziellem Datenschutz».