

# Anweisung

## Nutzung des allgemeinen Chats

### STOS – KI-Plattform

## Inhaltsverzeichnis

1. Dateneingabe.....	3
1.1 Personenbezogene und Personenbeziehbare Daten.....	3
1.2 Urheber- / Markenrecht.....	3
1.3 Besonders schützenswerte Daten .....	3
2. Datenausgabe.....	4
2.1 Personenbezogene Daten.....	4
2.2 Transparenz .....	4
2.3 Fehlerüberprüfung.....	5
2.4 Keine automatisierte Letztentscheidung.....	5
3. Datenschutz durch Technikgestaltung (Datenminimierung).....	5
3.1 Ausschluss der Nutzung zu Trainingszwecken.....	5
3.2 Transparenz und Wahlmöglichkeit hinsichtlich Eingabe-Historie .....	5
4. Freigegebene Nutzung .....	6
4.1 Beschreibung zulässiger Einsatzzwecke .....	6
4.2 Folgende Nutzung ist nicht freigegeben.....	6

## 1. Dateneingabe

Alle Mitarbeitenden, die das KI-System „Allgemeiner Chat“ der STOS KI-Plattform nutzen möchten, sind verpflichtet, die vorliegende Nutzungsrichtlinie zunächst sorgfältig zu lesen und die darin enthaltenen Vorgaben einzuhalten, um einen rechtskonformen Umgang mit diesem KI-System zu gewährleisten.

### 1.1 Personenbezogene und Personenbeziehbare Daten

Personenbezogene Daten (alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person beziehen, siehe Art. 4 DSGVO) dürfen nicht in die KI-Systeme eingegeben werden. Das betrifft jegliche Informationen, die Rückschlüsse auf Bürgerinnen/Bürger, Kundinnen/Kunden, Geschäftspartnerinnen/Geschäftspartner oder sonstige Dritte enthalten, und ebenso Daten der eigenen Beschäftigten.

Das gilt auch für solche Eingaben, die unter Umständen auf konkrete Personen bezogen werden können. Es reicht nicht, Namen und Anschriften aus der Eingabe zu entfernen. Auch aus dem Zusammenhang lassen sich gegebenenfalls Rückschlüsse auf Autorinnen/Autoren und Betroffene ziehen. Bei KI-Systemen, deren Bestimmung es ist, Querbezüge auch aus unstrukturierten Daten herzustellen, ist diese Gefahr besonders hoch.

*Beispiel: „Entwerfe ein Arbeitszeugnis im befriedigenden Bereich für einen Sachbearbeiter im Bürgeramt.“*

Die Eingabe kann Personenbezug aufweisen, wenn erkennbar ist, aus welchem Unternehmen sie zu welchem Zeitpunkt getätigt wurde.

### 1.2 Urheber- / Markenrecht

Urheber- oder markenrechtlich geschützte Daten sowie durch persönliche Bildrechte geschützte Daten dürfen nicht in die KI-Systeme eingegeben werden. Mitarbeitende müssen weiterhin sicherstellen, dass keine urheberrechtlich geschützten Ergebnisse ohne entsprechende Rechte oder Genehmigungen verwendet oder verarbeitet werden.

### 1.3 Besonders schützenswerte Daten

Sensible oder vertrauliche Daten oder Daten, die dem Amts-, Dienst- oder Betriebsgeheimnis unterliegen, dürfen nicht in KI-Systeme eingegeben werden. Zu den besonders schutzbedürftigen Daten gehören neben Sozialdaten, personenbezogene Daten gem. Art. 9 DSGVO, aus denen religiöse oder weltanschauliche Überzeugungen, eine Gewerkschaftszugehörigkeit oder die Präferenz für eine bestimmte politische Partei hervorgehen sowie Gesundheitsdaten und genetische oder biometrische Daten.

## 2. Datenausgabe

### 2.1 Personenbezogene Daten

Es ist darauf zu achten, dass verwendete Ergebnisse des KI-Systems keine personenbezogenen Daten enthalten.

Auch wenn der Eingabebefehl keine Person nennt, kann die KI unter Umständen vorherige Eingaben oder Informationen aus dem Internet einbeziehen. Daher sollten die Eingaben auf Fallgestaltungen beschränkt werden, die keinen Bezug zu Einzelpersonen herstellen.

*Beispiel einer unproblematischen Eingabe: „Schreibe einen SocialMedia Post zur Maiwoche 2025.“*

*Beispiel einer problematischen Eingabe: „Welche Personen der Stadtverwaltung waren auf der Maiwoche 2025?“*

### 2.2 Transparenz

In nachgelagerter Kommunikation ist hinreichende Transparenz über die Nutzung eines KI-Systems zu gewährleisten. Der Adressat ist, je nach Fall, darüber aufzuklären, dass KI zum Einsatz gekommen ist. Ggf. muss der Originaltext als Quelle angegeben werden. Es sind konkret zwei Fälle zu unterscheiden:

#### 1. Informationen von öffentlichem Interesse

Wer ein KI-System einsetzt, das Text generiert oder manipuliert, der zu dem Zweck veröffentlicht wird, die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse zu informieren, muss offenlegen, dass der Text künstlich generiert oder manipuliert wurde.

#### 2. Hilfsfunktionen für standardisierte redaktionelle Prozesse

Wer ein KI-System einsetzt, das einen Text generiert oder manipuliert, der den Zweck einer Hilfsfunktion für die Standardredaktion erfüllt oder die vom Anwender bereitgestellten Eingabedaten oder deren Semantik nicht wesentlich verändert, muss den Text nicht extra kennzeichnen.

*Beispiel einer Kennzeichnung: „Dieser Text wurde mit Unterstützung einer KI generiert.“*

*Beispiel einer Kennzeichnung: „Dieser Text wurde durch eine KI generiert.“*

### 2.3 Fehlerüberprüfung

Die Ausgabe muss auf Fehler (Halluzinationen)<sup>1</sup> geprüft werden und im Falle von Fehlern oder Ungenauigkeiten angepasst werden.

### 2.4 Keine automatisierte Letztentscheidung

Die unmittelbare Übernahme von KI-generierten Textvorschlägen in Entscheidungen mit Rechtswirkung ohne eine eigenständige inhaltliche Prüfung durch die zuständigen Mitarbeitenden ist unzulässig.

## 3. Datenschutz durch Technikgestaltung (Datenminimierung)

### 3.1 Ausschluss der Nutzung zu Trainingszwecken

Grundsätzlich sollen keine Ein- und Ausgabedaten für das Training verwendet werden, die Funktion ist entsprechend zu deaktivieren. Sollte ein Ausschluss der Nutzung zu Trainingszwecken nicht möglich sein und sind personenbezogene Daten betroffen, ist für diesen Zweck eine Rechtsgrundlage erforderlich. Datenschutzrechtlich vorzugswürdig sind daher Anwendungen, die die Ein- und Ausgabedaten nicht zu Trainingszwecken verwenden.

### 3.2 Transparenz und Wahlmöglichkeit hinsichtlich Eingabe-Historie

Viele durch Texteingaben (Prompts) gesteuerte Dienste bieten an, die Eingaben zu speichern, um z. B. den Dialog zu einem Thema zu einem späteren Zeitpunkt wieder aufnehmen zu können oder an einer weiteren Optimierung des Prompts zu arbeiten. Hierdurch wird eine Historie der Eingaben einer Person angelegt. Insbesondere bei der gemeinsamen Nutzung durch mehrere Beschäftigte muss dies transparent kommuniziert werden und die Möglichkeit für die Nutzer bestehen, selbst darüber zu entscheiden, ob die eigene Eingabe-Historie gespeichert wird.

---

<sup>1</sup> Als *Halluzinationen* wird in der KI der Umstand bezeichnet, dass Sprachmodelle Inhalte erzeugen, die zwar plausibel formuliert sind, aber nicht der Realität oder den zugrunde liegenden Daten entsprechen.

## 4. Freigegebene Nutzung

### 4.1 Beschreibung zulässiger Einsatzzwecke

Der Fokus der Anwendung liegt auf der kreativen Arbeit mit Texten, wie in den folgenden Beispielen:

- Stellenbeschreibungen
- Ideen für Projektnamen
- Textideen
- Texte in leichte bzw. einfache Sprache übersetzen
- Zusammenfassung öffentlicher Dokumente

Die Ergebnisse sollten immer als Entwurf bzw. Vorschlag behandelt und nicht 1:1 übernommen, sondern kritisch geprüft und angepasst werden.

### 4.2 Folgende Nutzung ist nicht freigegeben

Die Nutzung des allgemeinen Chats ist für folgende Zwecke nicht zulässig:

- Bilderstellung oder -manipulation
- Audioerstellung oder -manipulation
- Videoerstellung oder -manipulation

Bei oben genannten Zwecken ist das Risiko eines verwirrenden oder sogar täuschenden "Deep-Fakes" (realistisch wirkenden Medieninhalten, die durch KI-Systeme erzeugt oder manipuliert wurden) besonders hoch. Es gibt aktuell keine oder keine aufwandsarmen Erkennungssysteme für solche Inhalte und dementsprechende ausreichende Möglichkeiten die Transparenz vollumfänglich zu gewährleisten.

Daneben ist das Risiko versehentlich Urheber- oder Bildrechte zu verletzen entsprechend groß, so dass in Summe aktuell eine entsprechende Nutzung nicht freigegeben wird.