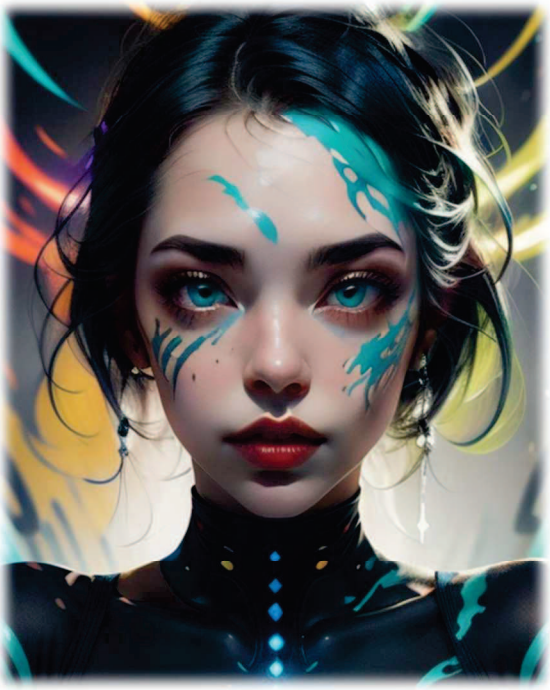
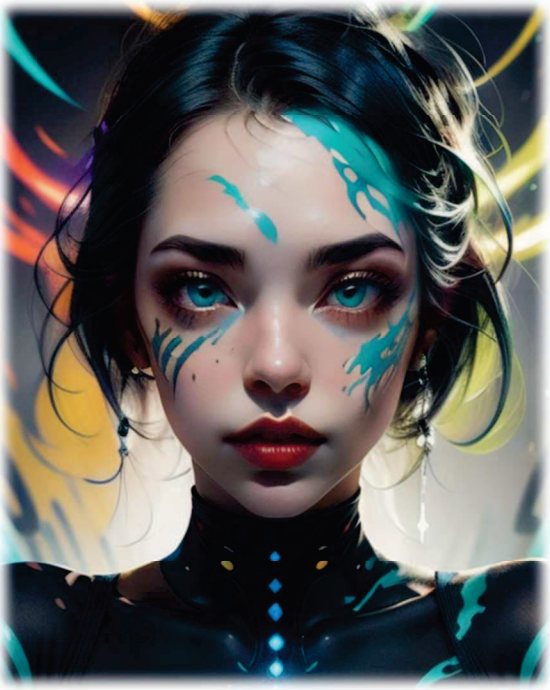

Datenschutzrisiken beim Einsatz frei verfügbarer Large Language Models (LLMs) durch Stadträte
Adobe Firefly



Warum ist das Thema relevant?

Large Language Models (LLMs) wie ChatGPT, Gemini, Claude oder frei verfügbare Open-Source-Modelle können die kommunale Arbeit, auch die von Gremienmitgliedern unterstützen, beispielsweise beim Erstellen von Textentwürfen, Zusammenfassen von Dokumenten, Recherchieren oder Strukturieren von Informationen.

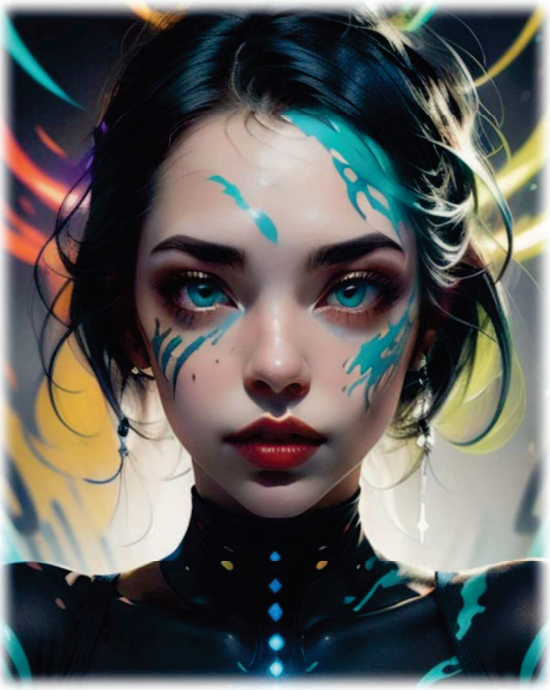
Die datenschutzrechtliche Herausforderung besteht darin, dass dabei unter Umständen schützenswerte Informationen an externe KI-Dienste übermittelt werden. Dazu können personenbezogene Daten von Bürgerinnen und Bürgern, Beschäftigten, Daten von Unternehmen oder politischen Mandatsträgern aber auch vertrauliche Verwaltungsinformationen gehören.



Risikofelder

Für eine Kommune sind insbesondere fünf Risikofelder zu betrachten:

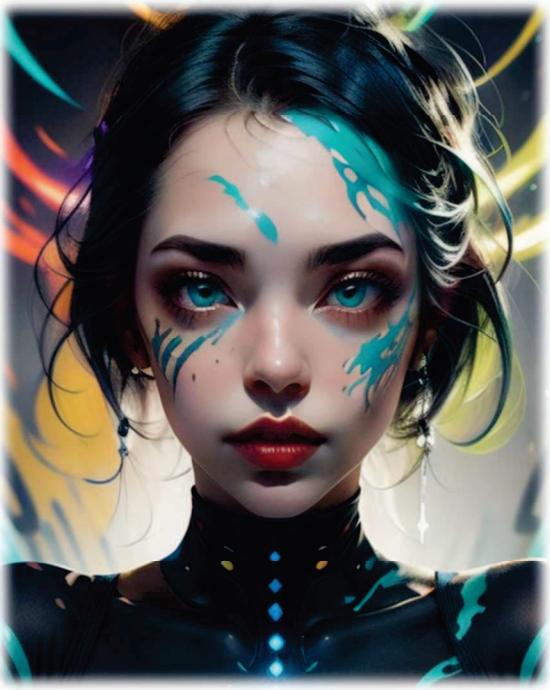
- Verarbeitung und mögliche Weiterverwendung eingegebener Daten im LLM und seinen ergänzenden Komponenten.
- Nutzung von Informationen zum Training von KI-Modellen.
- Profilbildung, Analyse und gegebenenfalls Nutzung zu Werbezwecken.
- Grundsätzliche Risiken externer, allgemein verfügbarer KI-Dienste, unter anderem mit Sitz außerhalb der EU.
- Zusätzliche Risiken, wenn mehrere Personen denselben Account verwenden.



Chatverlauf, Exportfunktionen und Restriktionen der Anbieter

Risiken

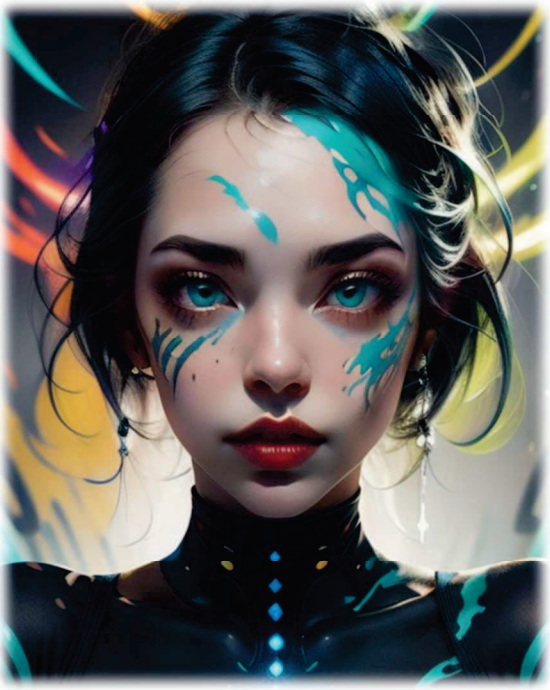
- Speicherung der Eingabe und der erzeugten Antwort im Chatverlauf macht die Daten dauerhaft verfügbar.
- Auswertung zur Erkennung von Missbrauch, Fehlern oder Sicherheitsproblemen durch die Anbieter.
- Überprüfung ausgewählter Gespräche durch Mitarbeitende der Anbieter.
- Verarbeitung durch eingebundene Dienstleister oder weitere technische Systeme.



Maschinelles Lernen

Eingaben werden in den Standardeinstellungen von LLMs fast immer zum fortlaufenden Training der Modelle des Anbieters genutzt.

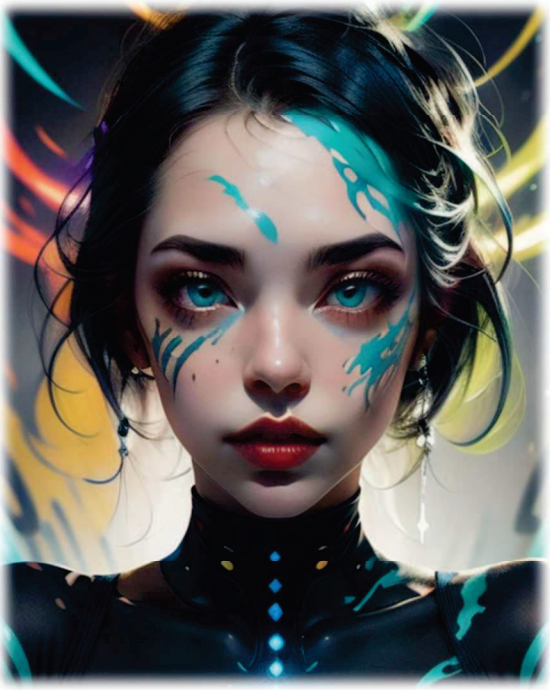
Die Verwendung der Eingaben zum maschinellen Lernen ist möglichweise abschaltbar. Bei kostenfreien Versionen ist dies bereits jetzt nicht mehr in jedem Falle möglich, Tendenz zunehmend. Dabei ergeben sich mehrere Problemfelder.



Maschinelles Lernen

Problemfelder

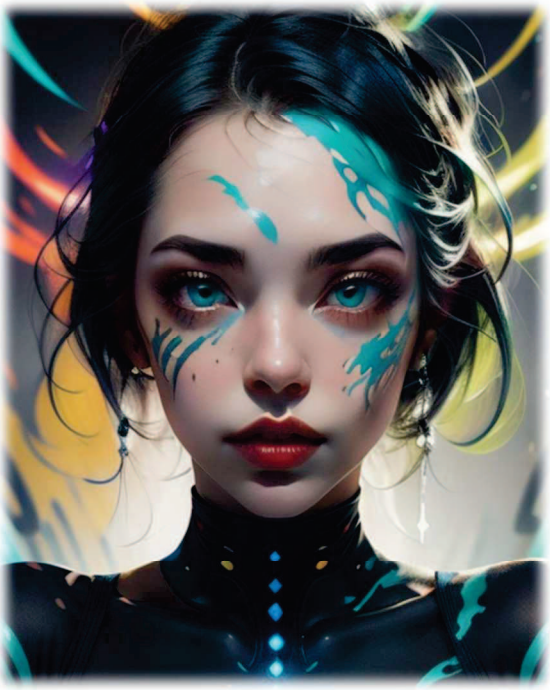
- Kontrollverlust: Es lässt sich nicht nachvollziehen, ob und wie eine Information in spätere Trainingsprozesse einfließt.
- Reproduktion: Ein Modell kann regelmäßig Inhalte aus Trainingsdaten wiedergeben. Das ist kein zwangsläufiges Ergebnis jeder Eingabe, aber ein bekanntes Risikofeld.
- Unmögliche Löschung: Die von einem öffentlichen LLM zum maschinellen Lernen verwendeten Informationen sind grundsätzlich nicht wieder löschbar.
- Zweckänderung: Informationen, die ursprünglich zur Bearbeitung eines bestimmten Vorgangs eingegeben wurden, könnten zu eigenen Zwecken durch den Anbieter des LLM verwendet werden.



Maschinelles Lernen – Ein Beispiel

Ein Ratsmitglied lädt einen Entscheidungsentwurf in ein LLM zur rechtlichen Bewertung. Der Entwurf enthält sowohl Angaben zu den Bearbeitern aus der Verwaltung wie auch von der Entscheidung betroffenen Dritten. Aus den Informationen lassen sich unter anderem Rückschlüsse auf die Bonität bzw. auf Geschäftsgeheimnisse ziehen.

- Für die Weitergabe der Daten an den Betreiber des LLM ist regelmäßig keine Rechtsgrundlage vorhanden.
- Auch wenn das Modell die Daten später nicht wörtlich reproduziert, können bei entsprechenden Nachfragen Teile davon reproduziert werden.

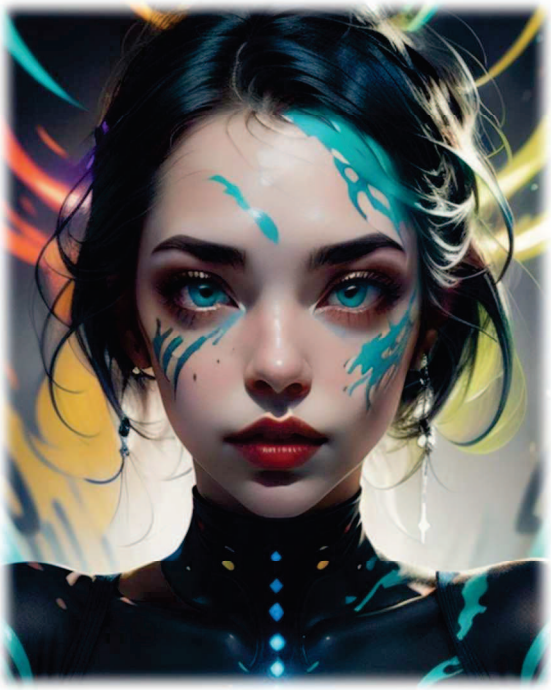


Nutzung von Informationen zu Werbezwecken und Profilbildung

Ein Anbieter kann im Rahmen seiner Dienste Informationen über die Nutzung erheben, beispielsweise:

- Welche Themen eine Person häufig bearbeitet.
- Welche Fragen sie stellt und welche Inhalte sie hochlädt.
- Welche Produkte, Dienstleistungen oder Interessen in ihren Gesprächen vorkommen.
- Zu welchen Zeiten und mit welchen Funktionen sie den Dienst nutzt.
- Welche weiteren Dienste oder Anwendungen mit dem Account verknüpft sind.

Mit den vorhandenen Informationen lassen sich regelmäßig Profile von Nutzern, möglicherweise auch von in den Daten genannten Personen erzeugen.

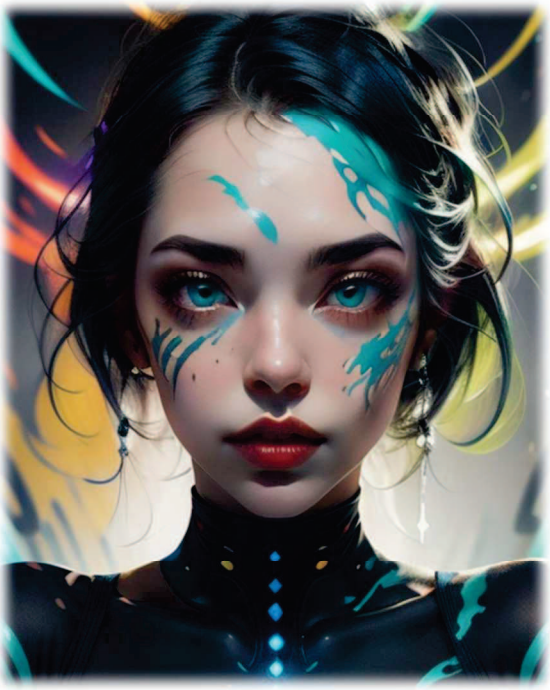


Nutzung von Informationen zu Werbezwecken und Profilbildung

All die erfassten Daten können zur Bildung von Profilen genutzt werden. Insbesondere bei durch Werbung finanzierten Modellen ist dies Standard.

In kommunalen Dokumenten finden sich regelmäßig Angaben, aus welchen sich Rückschlüsse auf besonders schützenswerte Lebensumstände oder Geschäftsgeheimnisse ableiten lassen.

- Einkommen, Sozialleistungen und finanzielle Notlagen.
- Personalprobleme und arbeitsrechtliche Konflikte.
- Bauvorhaben, Grundstücksangelegenheiten und wirtschaftliche Interessen.
- Politische Meinungen, interne Entscheidungsvorlagen und nicht veröffentlichte Vorhaben.

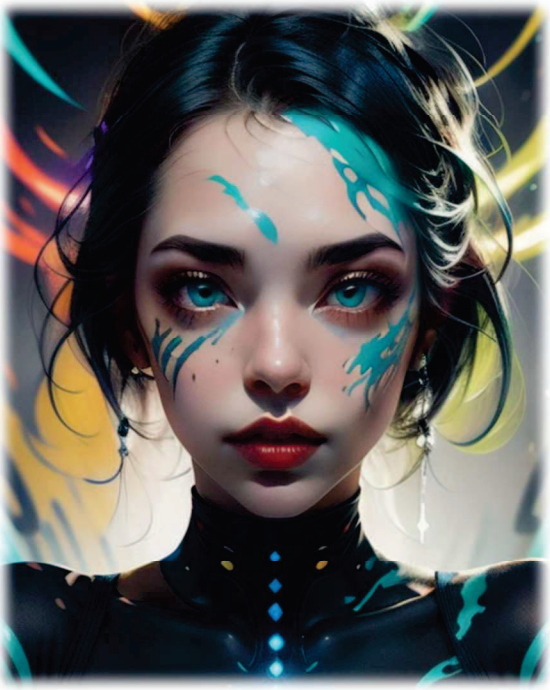


KI-Anbieter außerhalb der EU

Kernproblem

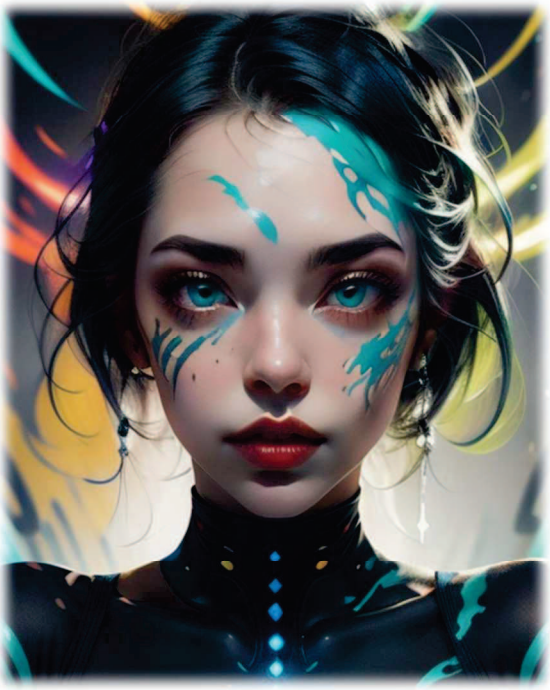
Werden Informationen aus der kommunalen Ratsarbeit durch einzelne Ratsmitglieder an einen KI-Anbieter außerhalb der EU übermittelt, wird für diese Art der Datenverarbeitung regelmäßig keine Rechtsgrundlage existieren. Besonders relevant ist dies bei nichtöffentlichen Sitzungsunterlagen, personenbezogenen Daten und vertraulichen Verwaltungsinformationen.

Lädt ein Stadtratsmitglied ohne Autorisierung Ratsinformationen in einen Dienst mit Sitz außerhalb der EU hoch, kann dies bereits per se eine Verletzung kommunaler Verschwiegenheitspflichten darstellen. Bei personenbezogenen Daten können sich zudem datenschutzrechtliche Folgen ergeben.



Grundlegende Risiken allgemein verfügbarer LLMs

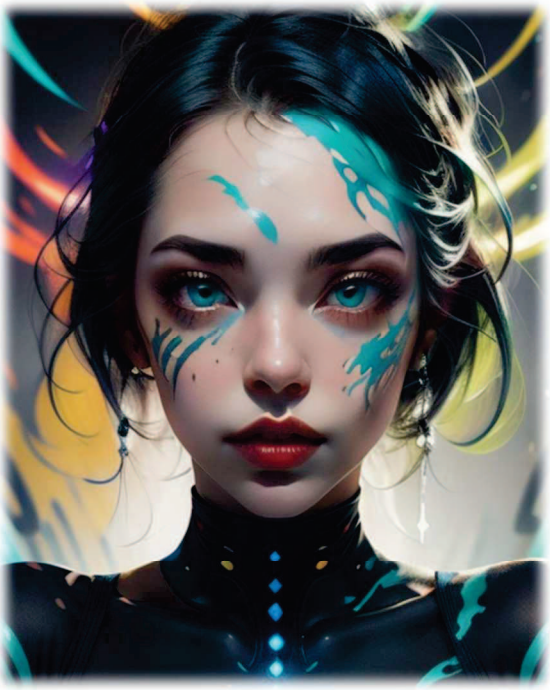
- Übermittlung personenbezogener Daten an externe Anbieter
- Speicherorte und mögliche Drittlandübermittlungen
- Unterauftragnehmer und Zugriffsrechte
- Fehlende Kontrolle über Speicherung und Löschung
- Unbeabsichtigte Offenlegung und Datenpannen
- Verknüpfung von Daten und Rückschlüsse auf Personen
- Unklare Verantwortlichkeit und eingeschränkte Nachvollziehbarkeit



Nutzung eines Accounts durch mehrere Anwender (Account-Sharing)

Aus der Nutzung eines KI-Accounts durch mehrere Benutzer ergeben sich hohe Datenschutz- und Informationssicherheitsrisiken, da sämtliche Informationen im Verlauf und ggf. weitere Meta-Daten für jeden Benutzer zur Verfügung stehen.

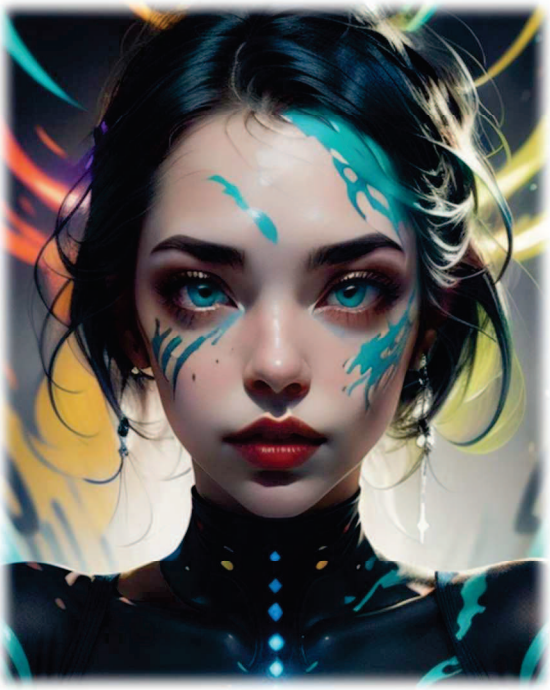
- Offenlegung aller vorhandenen Informationen und Dokumente.
- Nutzung und Weitergabe von gemeinsamen Passworten führt zu unkontrollierbaren Zugriffen.
- Eine Kompromittierung des Accounts betrifft sämtliche Nutzer.
- Bei einem Sicherheitsvorfall ist eine gezielte Sperrung oder Untersuchung einzelner Nutzeraktivitäten weitgehend unmöglich.



Vermischung beruflicher und privater Nutzung

Dadurch können sich unterschiedliche Datenbestände vermischen: private Interessen, berufliche Aufgaben, persönliche Notizen und kommunale Sachverhalte.

Je nach Funktionsumfang können auch gespeicherte Erinnerungen, individuelle Einstellungen oder verbundene Anwendungen die Nutzung beeinflussen. Anbieter könnten dadurch Informationen erhalten oder Kontext nutzen, der nicht für sie bestimmt war.



Rechtsfolgen bei unautorisiertem Einsatz von KIs

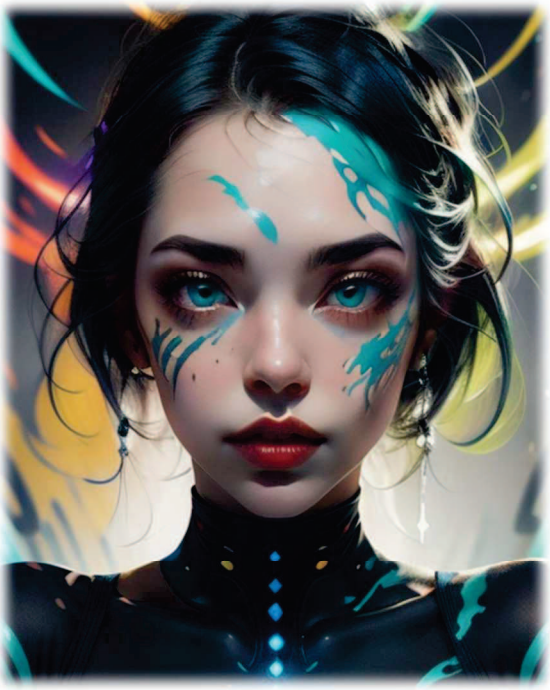
Ausgangsszenario

Ein Stadtratsmitglied lädt Informationen, die ihm aufgrund seines Mandats zugänglich sind, ohne Genehmigung in einen externen KI-Dienst hoch. Es wird angenommen, dass die Informationen nicht öffentlich zugänglich sind und personenbezogene oder vertrauliche Inhalte enthalten. Mögliche Rechtsfolgen sind:

Kommunalrecht: Ordnungsgeld bis 500 Euro

Nach § 19 Abs. 2 und Abs. 4 SächsGemO ist ein ehrenamtlich tätiges Stadtratsmitglied zur Verschwiegenheit über geheimhaltungsbedürftige Angelegenheiten verpflichtet. Es darf solche Informationen auch nicht unbefugt verwerten.

Verstößt es gegen diese Pflichten, kann der Stadtrat dem Mitglied ein Ordnungsgeld von bis zu 500 Euro auferlegen.



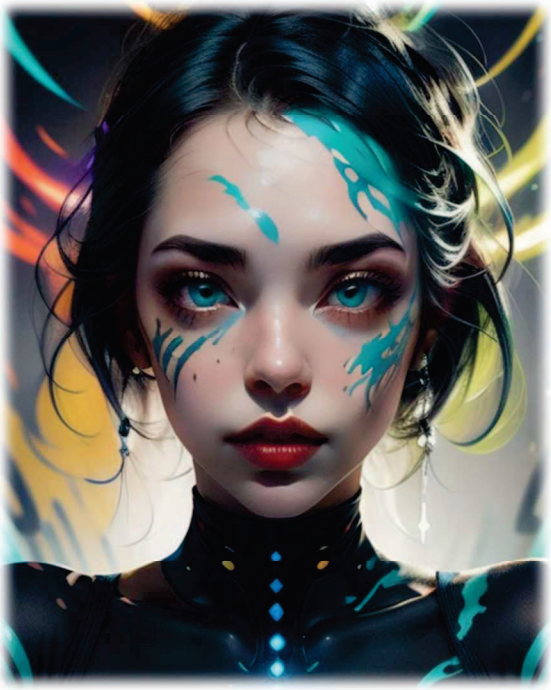
Rechtsfolgen bei unautorisiertem Einsatz von KIs

Persönlicher Schadensersatz und Haftung

Rechtsgrundlagen: Art. 82 DSGVO und § 823 BGB.

Wenn durch den Upload ein Datenschutzverstoß und ein nachweisbarer materieller oder immaterieller Schaden entstehen, kann ein Schadensersatzanspruch bestehen.

Der Anspruch nach Art. 82 DSGVO richtet sich grundsätzlich gegen den datenschutzrechtlich Verantwortlichen oder einen haftenden Auftragsverarbeiter. Bei einem Upload im Rahmen des Stadtratsmandats ist daher nicht automatisch das einzelne Ratsmitglied persönlich Schuldner. Eine persönliche zivilrechtliche Haftung nach § 823 BGB kann sich jedoch aus der Art der Verwendung der Daten und dem Vorgehen ergeben.



Rechtsfolgen bei unautorisiertem Einsatz von KIs

§ 353b StGB – Verletzung des Dienstgeheimnisses und einer besonderen Geheimhaltungspflicht

Die Vorschrift kann bei der unbefugten Offenbarung geschützter Geheimnisse und einer Gefährdung wichtiger öffentlicher Interessen einschlägig sein. Sie erfasst unter bestimmten Voraussetzungen auch fahrlässige Gefährdungen. Allerdings muss das Ratsmitglied zum gesetzlich erfassten Personenkreis gehören.

Beispiele für Datenschutzverletzungen

Weitergabe von Informationen an Dritte aus nichtöffentlichen Sitzungen

- Stadtratsmitglieder dürfen mit niemandem, der nicht selbst an der nichtöffentlichen Sitzung teilgenommen hat, über Inhalte sprechen. Das gilt auch für Familienmitglieder. Auch indirekte Informationen, aus welchen sich Schlüsse auf die Person oder die Situation ziehen lassen, fallen darunter.
  - Beispiel: Information über Zahlungsverzug, aus welcher sich Schlüsse auf die wirtschaftliche Situation eines Unternehmers ziehen lassen.

Bewerten lassen von Kosteninformationen oder Kalkulationen durch LLMs

- Unter Umständen können die Informationen aus dem LLM reproduziert werden. Bei aktiviertem Verlauf kann jeder Nutzer alle Informationen jederzeit nachvollziehen.

Übertragen oder Hochladen von Dokumenten der Verwaltung

- Verwaltungsdokumente enthalten regelmäßig persönliche Informationen zu Bearbeitern, Antragstellern oder weitere Angaben, aus welchen sich Personenbezüge ergeben. All diese Informationen landen unkontrollierbar bei den Anbietern der LLMs.



Quellen

Soweit nicht anderweitig angegeben stammen alle Bilder in dieser Präsentation von Pixabay
https://pixabay.com/

Selbsterstellte KI-Bilder wurden mit der Adobe KI Firefly generiert
https://firefly.adobe.com/