StK -Abteilung 3
Referat StK 30

[redacted] – CIO / StK 3
[redacted] - StK 30 Tel.: [redacted] Kiel, 02.05.2023

7860/2023

Zulässigkeit des Einsatzes von textbasierten Dialogsystemen unter Nutzung maschinellen Lernens in der Landesverwaltung Schleswig-Holstein
hier: Einsatz von ChatGPT des Unternehmens OpenAI

Vorschlag

Freigabeempfehlung für den Einsatz des textbasierten Dialogsystems ChatGPT des Dienstanbieters OpenAI (im Folgenden „Dienst“), gemäß § 2 Abs. 1 IT-Einsatz-Gesetz i.V.m. Ziffer 4.2 OrgErl ITSH durch den CIO Schleswig-Holstein unter Beachtung der im Übrigen geltenden gesetzliche Anforderungen des IT-Einsatz-Gesetzes mit den folgenden organisatorischen Maßgaben und Festlegungen:

1. ChatGPT des Anbieters OpenAI wird gemäß § 5 Abs. 1 IT-Einsatz-Gesetz als Assistenzsystem der Stufe 1 i.S.d. § 3 Abs. 2 IT-Einsatz-Gesetz eingestuft (Stand 14.03.2023).
2. Die finale Entscheidung über den Einsatz des Dienstes treffen die öffentlichen Stellen in eigener Verantwortung. Für die Beachtung der organisatorischen Vorgaben sowie die weiteren gesetzlichen Bestimmung bei der Verwendung des Dienstes ist die öffentliche Stelle verantwortlich, die den Dienst verwendet, § 4 Abs. 1 IT-Einsatz-Gesetz. Dies gilt auch für die Sicherstellung der sonstigen Rechtmäßigkeit des Einsatzes und die Einhaltung der bei der Verwendung geltenden Nutzungsbedingungen des Anbieters OpenAI.
3. Die Verwendung von

   (a) im öffentlichen Interesse als geheimhaltungsbedürftig eingestuften Informationen i.S.d. § 5 und § 2 Abs. 2 Landessicherheitsüberprüfungsgesetz sowie

   (b) Informationen, die gemäß § 9 und § 10 IZG SH, insbesondere personenbezogenen Daten i.S.d. Art. 4 Abs. 1 DSGVO, oder entsprechenden Regelungen in den Fachgesetzen nicht veröffentlicht werden dürften,

   zur Erstellung einer Abfrage (Input) ist unzulässig.
4. Durch die nutzende Person bzw. öffentliche Stelle muss sichergestellt sein, dass die für die Abfrage genutzten Informationen oder Daten frei von Rechten Dritter

sind, siehe dazu auch Ziffer 3 (a) Terms of Use (Stand 14.03.2023 – https://openai.com/policies/terms-of-use).

5. Die Nutzung des Dienstes unter Verwendung personenbezogener Daten zum Zweck der Erstellung von Beurteilungen der Persönlichkeit, der Arbeitsleistung, der physischen und psychischen Belastbarkeit, der kognitiven oder emotionalen Fähigkeiten von Menschen, der Erstellung von Prognosen über die Straffälligkeit einzelner Personen oder Personengruppen ist gemäß § 2 Abs. 2 Nr. 2 IT-Einsatz-Gesetz nicht zulässig.
6. Die unmittelbare Verwendung der durch den Dienst erstellten Textbausteine (Output) zum Zweck des Erlasses eines Verwaltungsakts ist gemäß § 2 Abs. 2 IT-Einsatz-Gesetz unzulässig.
7. Werden die Ergebnisse des Dienstes (Output) ganz oder teilweise zur Erstellung von Texten im dienstlichen Kontext, z.B. zur Generierung von Textbeiträgen, Bausteinen für Reden und Antworttexten, Begründung von Entscheidungen oder Vermerken verwendet, ist dies gemäß § 6 Abs. 4 IT-Einsatz-Gesetz durch die Formulierung:

   „Unter Verwendung des textbasierten Assistenzsystems ChatGPT - https://chat.openai.com/ (Automationsstufe 1 gemäß §5 IT-Einsatz-Gesetz) erstellt.“

   kenntlich zu machen.
8. Die Abfrage (Input) ist zusammen mit dem Ergebnis (Output) Teil des aktenrelevanten Handelns der öffentlichen Stelle, zusammen mit dem Gesamtvorgang zu den Akten zu nehmen, § 8 Abs. 4 IT-Einsatz-Gesetz.
9. Der Dienst kann ungenaue Informationen über Personen, Orte oder Fakten erstellen, daher sind Beschäftigte der öffentlichen Stellen verpflichtet, die fachliche und sachliche Korrektheit der erstellten Ergebnisse über eine unabhängige Quelle vor der Verwendung zu dienstlichen Zwecken zu überprüfen. Die gilt auch für diskriminierende oder dem deutschen Recht widersprechende Ergebnisse.
10. Öffentliche Stellen des Landes müssen in eigener Verantwortung sicherstellen, dass der Schutz der gegenüber OpenAI zur Nutzung des Dienstes ggfs. übermittelten personenbezogenen Daten insbesondere Accountinformationen der Beschäftigten beachtet wird, z.B. in dem Funktionsaccounts angelegt werden und für die Nutzung des Dienstes keine privaten Accountdaten verwendet werden.

Anlass

Mit Schreiben vom 21.04.2023 hat das MWVATT StK 3 gebeten den Einsatz von ChatGPT für dienstliche Zwecke zu prüfen. Aus dem Wirtschaftsministerium heraus, welches auch für Technologie und Innovation zuständig ist, kommt zunehmend die Bitte ChatGPT bzw. eine andere Software der künstlichen Intelligenz für Recherchen, Generierung von Textbeiträgen, Bausteine für Reden und Antworttexte u.a. unentgeltlich zu nutzen. Das MWVATT befürwortet die Verwendung von ChatGPT als zusätzliches Hilfsmittel.

Sachstand

ChatGPT ist ein textbasiertes Dialogsystems, welches über eine Benutzerschnittstelle Eingaben (Input) entgegen nimmt und dann unter Verwendung maschinellen Lernens eine Ausgabe (Output) generiert (siehe https://de.wikipedia.org/wiki/ChatGPT und https://platform.openai.com/docs/models/overview). Je nach verwendetem Modell, kann ein unterschiedlicher Reifegrad der textbasierten Interaktion erreicht bzw. genutzt werden. So kann das Modell unter Verwendung von breitem Allgemeinwissen und Fachkenntnissen auf komplexe Anweisungen in natürlicher Sprache reagieren und sowohl natürliche Sprachausgaben als auch Code generieren.

Stand 3. Mai 2023 wird in der kostenfreien Variante von ChatGPT das Sprachmodell GPT-3.5 eingesetzt. Die Veröffentlichung der Firma OpenAI lassen vermuten, dass ab 10. Mai 2023 das Modell GPT-4 auf für die kostenfreie Variante zum Einsatz kommt.

Der Dienst wird durch das Unternehmen OpenAI, mit Sitz in San Francisco, USA entwickelt und betrieben:

OpenAI, L.L.C.
3180 18th St
San Francisco, CA 94110

1. Allgemeine Zulässigkeit

   Nach dem verfassungsrechtlichen Prinzip der Gesetzmäßigkeit der Verwaltung und daraus fließenden Prinzip des Vorbehalts des Gesetzes gemäß Art. 20 Abs. 3 GG bedarf das Handeln der öffentlichen Verwaltung insbesondere dann einer gesetzlichen Grundlage, wenn dadurch wesentliche Fragen des Staates oder des Gemeinwohls betroffen sind. Der teilweise oder vollständige Einsatz maschinellen Lernens anstelle menschlicher Handlungen in Ausübung öffentlicher Gewalt, ist aus ganz grundsätzlichen Erwägungen heraus eine wesentliche Entscheidung. Der Einsatz algorithmusgesteuerter Entscheidungsfindung nicht allein im öffentlichen Bereich ändert sehr grundsätzlich das Verhältnis zwischen Staat und Gesellschaft und erfordert teilweise neue Methoden und Maßnahmen zur Gewährleistung der verfassungsrechtlich normierten Rechtsordnung, der staatlichen Organisation, dem demokratischen und rechtsstaatlichen Ausgleich widerstreitender Interessen und dem Schutz und der Gewährung der Grund- und Menschenrechte.
   Der Einsatz von ChatGPT stellt hier, auch wenn die tatsächlichen Auswirkungen im Hinblick auf die Bandbreite des Handelns der schleswig-holsteinischen Verwaltung derzeit nur minimal sind, den Beginn eines grundsätzlichen Wandels der Entscheidungsfindung und staatlichen Aufgabenerledigung dar. Es handelt sich hierbei um eine wesentliche Änderung in der Ablauforganisation und dem Einsatz von Informationstechnologie in der öffentlichen Verwaltung des Landes Schleswig-Holstein.
   Daher bedarf es auch für den Einsatz eines dialogbasierten Assistenzsystem wie ChatGPT einer entsprechenden gesetzlichen Grundlage. Dies hat der Gesetzgeber mit dem *Gesetz über die Möglichkeit des Einsatzes von datengetriebenen Informationstechnologien bei öffentlich-rechtlicher Verwaltungstätigkeit (IT-Einsatz-Gesetz – ITEG) vom 16. März 2022* geschaffen. § 2 Abs. 1 IT-Einsatz-Gesetz erlaubt den Einsatz dieses Dienstes.
   Zum Schutz der entsprechenden öffentlichen und privaten Interessen sowie der Gewährleistung eines fachlich und inhaltlich korrekten sowie rechtlich sicheren Einsatzes des Dienstes und Gewährleistung der Prinzipien der Transparenz, Beherrschbarkeit, Robustheit und Sicherheit müssen, gemäß der mit dem Einsatz verbundenen Risiken, technische und/oder organisatorische Maßnahmen durch die

verwendende Stelle getroffen werden, § 2 Abs. 1 Satz 2 IT-Einsatz-Gesetz. Diese sind allerdings abhängig von dem durch die genutzte Technologie verursachten Risiko, welches gemäß § 3 Abs. 2 IT-Einsatz-Gesetz auf der Grundlage der Automationsstufe ermittelt wird. Daraus leiten sich dann die zu ergreifenden organisatorischen und ggfs. technischen Maßnahmen ab.

2. Zu Ziffer 1 – Einstufung ChatGPT

   ChatGPT (GPT-3 und GPT-4) ist ein System welches durch Eingabe von Text selbständig aus verschiedenen Datenquellen eine Ausgabe, in der Regel als zusammenhängenden Text, erzeugt. Es handelt sich insoweit um eine Technologie, die für eine ihr zugewiesene Aufgabe selbstständig die Auswahl der relevanten Informationen sowie eine Priorisierung entscheidungsrelevanter Faktoren vornimmt. Die bearbeitende Person entscheidet nach der Generierung des Output, ob das erstellte Ergebnis für die zu erledigende Aufgabe eingesetzt und ob das gelieferte Ergebnis angenommen, abgelehnt oder unter Verwendung neuer Parameter wiederholt zur Bearbeitung gestellt wird. Die Einstufung erfolgt daher als Assistenzsystem (Stufe 1) gemäß § 3 Abs. 2 Nr. 1 IT-Einsatz-Gesetz.
   Auf der Grundlage dieser Einstufung leiten sich die zu ergreifenden organisatorischen Maßnahmen ab. Die Anordnung technischer Maßnahmen vor diesem Hintergrund und dem derzeitigen Kenntnisstand ist nicht erforderlich.

3. Zu Ziffer 2 – Verantwortlichkeit
   ChatGPT kann nur als externer Dienst über die entsprechend angebotenen Schnittstellen genutzt werden. Ein Einsatz als „on premise“ Lösung ist derzeit nicht vorgesehen auch ist die zentrale Steuerung des Dienstes und deren Einsatz nicht realisierbar.Der Dienst kann damit zur Zeit auch nicht als Basisdienst i.S.d. E-Government-Gesetzes eingestuft werden.

   Die endgültige fachliche, rechtliche und organisatorische Verantwortung für die Nutzung des Dienstes wird daher nicht zentral durch das ZIT übernommen. Die finale Einsatzentscheidung und die daraus entspringenden rechtlichen Konsequenzen tragen gemäß § 4 Abs. 1 IT-Einsatz-Gesetz.
   Die im Folgenden angeordneten organisatorischen Maßnahmen erfüllen dabei den Mindeststandard der gemäß § 4 Abs. 3 IT-Einsatz-Gesetz zu ergreifenden Maßnahmen. Die jeweiligen Einsatzstellen, können darüber hinaus weitere organisatorische Maßnahmen ergreifen.

4. Zu Ziffer 3 und 4 – Begrenzung der Verwendung von Informationen für den Input
   In der Standardeinstellungen nutzt ChatGPT auch die Eingaben der Nutzenden, um die Entscheidungsfindung zu verbessern. Daher muss davon ausgegangen werden, das alle an den Dienst übermittelten Informationen zu einem späteren Zeitpunkt anderen Nutzern gegenüber zur Generierung des Output verwendet werden (siehe Ziffer 3 Terms of Use). Derzeit sollte diese Annahme zur Grundlage der weiteren Entscheidungen über den Einsatz des Dienstes gemacht werden, auch wenn OpenAI anbietet, dass Informationen, die durch die Nutzenden zur Verfügung gestellt werden, (Non-API Content) von der Verbesserung des Modells ausgeschlossen werden können (siehe https://help.openai.com/en/articles/5722486-how-your-data-is-used-to-improve-model-performance).

   Zur Generierung einer Eingabe dürfen daher keine geheimhaltungsbedürftigen oder sicherheitsrelevanten Informationen genutzt werden. Dies gilt auch für Betriebs- und Geschäftsgeheimnisse und urheberrechtlich geschützte Informationen an denen das Land Schleswig-Holstein keine entsprechenden Nutzungsrechte besitzt.

Zudem ist eine Verwendung von Informationen untersagt, für die aufgrund fachgesetzlicher Regelungen ein Verbot der Veröffentlichung existiert.

Bezüglich personenbezogener Daten wird derzeit der Ausschluss der Verwendung empfohlen.
Grundsätzlich erscheint die Zulässigkeit der Nutzung personenbezogener Daten durch entsprechende rechtliche und organisatorische Maßnahmen realisierbar. So bietet OpenAI den Abschluss eines entsprechenden Auftragsverarbeitungsvertrages an (https://openai.com/policies/data-processing-addendum). Da es sich hierbei allerdings um eine Übermittlung staatlicherseits erhobener Daten in das nichteuropäische Ausland handelt, müssten Maßnahmen des Kapitel 5 der DSGVO ergriffen werden.
Nach hiesiger Einschätzung, würde dies zu einem im Hinblick auf den Nutzen des Einsatzes des Dienstes unverhältnismäßigen Aufwand seitens der nutzenden Stellen führen.

5. Zu 5 und 6 – Nutzungsbeschränkungen
   Die Beschränkung der Verwendung des Dienstes zu den genannten Zwecken beruht auf der gesetzlichen Nutzungsbeschränkung des § 2 Abs. 2 IT-Einsatz-Gesetz. Zu beachten ist hierbei allerdings, dass die Nutzung des Dienstes zur Unterstützung der Begründung eines zu erlassenden Verwaltungsaktes zulässig ist. Lediglich die Nutzung des Outputs z.B. auf die Frage, ob z.B. ein Gebäude abgerissen werden darf etc. ohne entsprechende rechtliche und fachliche Prüfung, wäre rechtswidrig.

6. Zu Ziffer 7 und 8 – Transparenz
   Eine der maßgeblichen Forderungen bei der Nutzung von Assistenzsystemen und Künstlicher Intelligenz ist die Sorge um die fehlende Transparenz. Mit dieser Maßnahme wird diesem gesetzlich verankerten Prinzip Rechnung getragen. Die Anforderung des § 6 Abs. 1 und Abs. 4 IT-Einsatz-Gesetz bezüglich der Offenlegung des Modells bzw. des Algorithmus kann mit Verweis auf die von OpenAI zur Verfügung gestellten Dokumentation unter https://platform.openai.com/docs/introduction begegnet werden.

   Es besteht die Möglichkeit, dass in den Antworttexten von ChatGPT eine Art digitales Wasserzeichen (z.B. charakteristische Abfolgen von Buchstaben über mehrere Wörter hinweg) enthalten ist, mit dem nachgewiesen werden kann, dass der Text mittels ChatGPT erstellt wurde. Eine Kennzeichnung der mit Hilfe von ChatGPT erzeugten Texte schafft Transparenz und verhindert eine „Enthüllung“ durch Dritte, dass in der Landesverwaltung heimlich ChatGPT eingesetzt wird.

   Die Ergreifung weiterer gesetzlich vorgesehener, ggfs. zentral zu realisierender Transparenzmaßnahmen, wie z.B. einem entsprechenden Register, werden derzeit durch StK 30 geprüft.

   Die aktenmäßige Dokumentation der Verwendung dient neben den Anforderungen an die Transparenz auch der Gewährleistung der Aktenmäßigkeit des staatlichen Handelns sowie der Ermöglichung der Revision und Prüfung der Zweckmäßigkeit des Einsatzes des Dienstes. Die Dokumentation von Input und Output kann z.B. als einfache Textdatei in dem entsprechenden Vorgang abgelegt werden.

7. Zu 9 – Inhaltskontrolle

   OpenAI weist selbst explizit darauf hin, dass die Ergebnisse (Output) fehlerhaft, diskriminierend oder beleidigend sein können. Letztlich kann damit auch nicht ausgeschlossen werden, dass durch die ungefilterte Nutzung des Dienstes,

Rechtsverstöße begangen werden können, z.B. Beleidigungen oder Verleumdungen.

Mit der Pflicht zur nachträglichen Überprüfung der Nutzung der Ergebnisse als verbindliche organisatorische Maßnahme, wird damit eine eigentlich als Selbstverständlichkeit anzusehende Anforderung, zu einer verbindlichen Maßnahme. Diese setzt zudem das Prinzip der Menschlichen Aufsicht, des Vorrangs menschlicher Entscheidungen und die Möglichkeit der Korrektur der maschinellen Entscheidung gemäß § 7 IT-Einsatz-Gesetz um.

8. Zu 10 – Accountinformationen
   Der Dienst kann nur mit einem entsprechenden Konto genutzt werden. Die Verwendung privater Accounts (z.B. von Google) im dienstlichen Kontext würde zu der auch bei der Nutzung von personenbezogenen Daten für den Input beschriebenen rechtlichen Situation führen. Die Empfehlung ist daher, den Dienst nur über dienstliche Geräte und mit einem dafür eigens angelegten ggfs. durch mehrere Personen genutzten Nutzerkontos zu nutzen.

Bewertung

1. Der Einsatz von ChatGPT ist unter den als verhältnismäßig und praxisnah einzuordnenden organisatorischen Maßnahmen rechtssicher möglich. Wesentliche Beschränkungen existieren nicht.
2. Die Entscheidung sollte mit einer entsprechenden Information über das Kabinett kommuniziert werden.
3. Sinnvoll wäre, in ca. 6 Monaten eine verbindliche Abfrage über die IMAG Digitalisierung durchzuführen, um einen Überblick über die Erfahrungen mit dem Einsatz des Dienstes zu bekommen. In diesem Zusammenhang sollte auch der weitere Bedarf an fachspezifischen Anwendungen von Large Language Models (LLM) wie ChatGPT diskutiert und ggfs. projekthaft auf Basis von data[port]ai realisiert werden.
4. Ggfs. steht diese Entscheidung im Widerspruch zu bisherigen, fachspezifischen Aussagen bzgl. der Zulässigkeit des Einsatzes von ChatGPT durch die Ressorts. Hier wäre eine entsprechende erneute rechtliche Prüfung empfehlenswert, da evtl. die Regelungen des IT-Einsatzgesetzes nicht vollständig in die Prüfung eingeflossen sind.

[redacted] (StK 30)