

Landesbeauftragte für Datenschutz · Postfach 71 16 · 24171 Kiel

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

Vorab per E-Mail an

Landesbeauftragte für Datenschutz

Holstenstraße 98

24103 Kiel

Tel.: 0431 988-1200

Fax: 0431 988-1223

Ansprechpartner/in:

Durchwahl: 988-1200

Aktenzeichen:

LD4-90.41/23.017

Kiel, 19. April 2023

Important Notice:

Official Request from the Data Protection Supervisory Authority of the German Federal State of Schleswig-Holstein

Anhörung nach § 87 Abs. 1 des Landesverwaltungsgesetzes Schleswig-Holstein (LVwG) im Rahmen der Aufgabenwahrnehmung nach Art. 57 Abs. 1 lit. a) der Verordnung (EU) 2016/679 (Datenschutz-Grundverordnung – DSGVO)

hier: Verfahren zur Prüfung des Dienstes ChatGPT und der zugehörigen Sprachmodelle GPT bis GPT-4

Sehr geehrte Damen und Herren,

die Landesbeauftragte für Datenschutz ist die Datenschutzaufsichtsbehörde des deutschen Bundeslandes Schleswig-Holstein. Sie überwacht gemäß Art. 57 Abs. 1 lit. a) der Datenschutz-Grundverordnung (DS-GVO), § 40 Abs. 1 des Bundesdatenschutzgesetzes (BDSG) und § 17 Abs. 1 Satz 3 des Landesdatenschutzgesetzes Schleswig-Holstein die Anwendung, Einhaltung und Durchsetzung der Vorschriften über den Datenschutz in Schleswig-Holstein.

Der von OpenAI, L.L.C. angebotene Dienst ChatGPT und die zugehörigen Sprachmodelle GPT bis GPT-4 basieren unter anderem auf der automatisierten Verarbeitung personenbezogener Daten. Da Sie die Nutzung des Dienstes auch Personen in der Europäischen Union anbieten, findet die Datenschutz-Grundverordnung (DS-GVO) gemäß deren Art. 2 Abs. 1 und Art. 3 Abs. 2 auf die Verarbeitung von personenbezogenen Daten mittels des Dienstes Anwendung.

Ihr Unternehmen unterhält keine Niederlassung in der Europäischen Union. Somit sind alle europäischen Aufsichtsbehörden für Datenschutz gemäß Art. 55 Abs. 1 DS-GVO im Hoheitsgebiet ihres jeweiligen Mitgliedsstaates zuständig für die Überwachung der Einhaltung der DS-GVO und die Ausübung der darin geregelten Befugnisse. Da der Dienst von Ihnen auch in der Bundesrepublik Deutschland und im Bundesland Schleswig-Holstein angeboten wird, ist meine Zuständigkeit begründet.

Es bestehen grundsätzliche Bedenken, ob die Verarbeitung von personenbezogenen Daten im Rahmen von ChatGPT den Anforderungen des europäischen und des deutschen Datenschutzrechts genügt. So erscheint es insbesondere fraglich, ob die Datenverarbeitung den in Art. 5 und Art. 25 DS-GVO festgeschriebenen, datenschutzrechtlichen Grundprinzipien gerecht wird, ob sie auf einer gültigen Rechtsgrundlage nach Art. 6 DS-GVO beruht und ob die spezifischen Anforderungen des Art. 8 an die Datenverarbeitung von Minderjährigen sowie des Art. 13 und 14 DS-GVO an die transparente Information der betroffenen Personen eingehalten werden.

Dies betrifft insbesondere die Erhebung der personenbezogenen Daten aus öffentlich zugänglichen Quellen, die Nutzung dieser personenbezogenen Daten als Trainingsdaten im Rahmen des maschinellen Lernens, die Speicherung der personenbezogenen Daten infolge des maschinellen Lernens und die Verarbeitung der personenbezogenen Daten der Nutzenden von ChatGPT. Wir bitten Sie, soweit möglich, die unten gestellten Fragen auf diese Verarbeitungsschritte zu beziehen und gesondert zu beantworten.

Um dies überprüfen zu können, benötige ich weitergehende Informationen und Unterlagen von Ihnen. Art. 31 DS-GVO sieht vor, dass der Verantwortliche mit der Aufsichtsbehörde bei der Erfüllung ihrer Aufgaben zusammenarbeitet. Die Ihnen vorliegende Anhörung dient dazu, über weitere Maßnahmen nach Art. 58 Abs. 1 und 2 DS-GVO, einschließlich des Erlasses einer Anordnung auf Erteilung einer Auskunft und der Bereitstellung von Unterlagen, zu entscheiden. Sie erhalten nach § 87 Abs. 1 des Landesverwaltungsgesetzes Schleswig-Holstein Gelegenheit, sich zu den für die Entscheidung maßgeblichen Tatsachen zu äußern.

Ich bitte Sie daher um die Beantwortung der folgenden Fragen sowie um die Vorlage jeweils ergänzender Unterlagen:

- I. Grundsätze der Datenverarbeitung
 - a. Wie stellen Sie sicher, dass bei ChatGPT alle Grundsätze für die Verarbeitung personenbezogener Daten aus Art. 5 Abs. 1 DS-GVO eingehalten werden?
 - b. Wie werden die Grundsätze der Zweckbindung, der Datenminimierung und der Speicherbegrenzung (Art. 5 Abs. 1 lit. b) 1. Hs., lit. c) und lit. e) DS-GVO), insbesondere hinsichtlich der für das Training der KI verwendeten personenbezogenen Daten, sichergestellt, wonach nur die für die jeweilige Funktionalität erforderlichen Daten erhoben und diese nur so lange gespeichert werden dürfen, wie es für die Zwecke, für die sie verarbeitet werden, erforderlich ist?
 - c. Wie lange speichern Sie die personenbezogenen Daten?
 - d. Auf welche Weise wird die Richtigkeit der verwendeten Daten und der personenbezogenen Daten, die ChatGPT daraus erzeugt, (Art. 5 Abs. 1 lit. d) DS-GVO), sichergestellt, insbesondere wenn eine betroffene Person das Recht auf Berichtigung (Art. 16 DS-GVO) oder das Recht auf Löschung (Art. 17 DS-GVO) in Anspruch nimmt?

- e. Bitte weisen Sie die Einhaltung dieser Grundsätze durch geeignete Dokumentation gemäß Art. 5 Abs. 2 DS-GVO nach.

II. Rechtmäßigkeit der Datenverarbeitung

- a. Aus welchen Quellen (konkrete Bezeichnung von natürlichen und juristischen Personen, Einrichtungen oder anderen Stellen; Firmenbezeichnungen von Unternehmen und Behörden; sämtliche Stellen, welche als Verantwortliche oder Auftragsverarbeiter in Betracht kommen – Art. 4 Nr. 7 und 8 DS-GVO) werden die Daten erhoben, mit denen der Dienst agiert und anhand derer die KI, insbesondere GPT-3 und GPT-4, trainiert wird?
- b. Auf welche Rechtsgrundlagen werden die verschiedenen Verarbeitungsvorgänge von personenbezogenen Daten bei ChatGPT jeweils gestützt?
 - i. Sofern die jeweilige Verarbeitung auf eine Einwilligung (Art. 6 Abs. 1 UAbs. 1 lit. a) und Art. 7 DS-GVO) der Nutzenden gestützt wird, legen Sie bitte das Muster der Einwilligungserklärung vor.
 - ii. Soweit Art. 6 Abs. 1 UAbs. 1 lit. f) DS-GVO als Rechtsgrundlage herangezogen wird, erläutern Sie bitte die Hintergründe und Erwägungen der jeweiligen Interessenabwägung.
- c. Wird bei der Erhebung und Weiterverarbeitung danach differenziert, ob die genutzten Daten Personenbezug aufweisen oder nicht?
- d. Werden von Nutzenden erhobene und eingegebene Daten neben der Generierung von Antworten zusätzlich auch als Trainingsdaten für Methoden des maschinellen Lernens genutzt?
- e. Werden ausdrückliche Einwilligungen der betroffenen Personen eingeholt, wenn besonders schützenswerte Daten nach Art. 9 Abs. 1 DS-GVO verarbeitet werden?
- f. Wenn Sie die Verarbeitung besonders schützenswerter Daten mit Art. 9 Abs. 2 lit. e) DS-GVO begründen, legen Sie uns bitte Ihre Bewertung der Voraussetzungen dar.
- g. Werden durch Schlussfolgerungen oder Kombination von Daten besonders schützenswerte Daten nach Art. 9 Abs. 1 DS-GVO abgeleitet? Wenn ja, auf welche Rechtsgrundlage stützen Sie eine solche Verarbeitung?
- h. Gibt es Maßnahmen, um Schlussfolgerungen oder Kombinationen von besonders schützenswerten Daten und deren weitere Verarbeitung zu verhindern, auch wenn eine solche Verarbeitung nicht gezielt erfolgt? Dies gilt sowohl für OpenAI als auch für die Nutzung durch andere Dienste oder Unternehmen.
- i. Findet Profilbildung von Nutzenden des Dienstes (Tracking) statt?
- j. Falls ja, zu welchen Zwecken werden die Profile genutzt? Dienen so erhobene Daten auch zu Zwecken der Werbung oder des Trainings für Methoden des maschinellen Lernens?

- k. Werden Analysedaten zur Nutzung des Dienstes erhoben und, falls ja, zu welchen Zwecken, auf welche Weise und in welchem Umfang? Was ist insbesondere unter den Begriffen „usage data“, „analytics“ sowie „variety of online analytic products“, welche in der Privacy Policy von OpenAI genannt werden, zu verstehen?
- l. Weshalb müssen Nutzende neben der E-Mail-Adresse auch die Telefonnummer eingeben? Zu welchen Zwecken wird die erhobene Telefonnummer verwendet? Wie lange wird die Telefonnummer durch OpenAI als Verantwortlichen gespeichert?

III. Rechte betroffener Personen, Transparenz

- a. Wie wird die Beachtung der Rechte betroffener Personen aus Art. 12 ff. DSGVO sichergestellt (insb. Rechte auf Transparenz, Auskunft, Berichtigung, Löschung)?
- b. Wie und an welcher Stelle werden die Nutzenden über die Erhebung und Verwendung personenbezogener Daten im Sinne von Art. 13 und 14 DSGVO informiert? Bitte fügen Sie Ihrer Antwort die entsprechenden Texte bei.
- c. Wird über die involvierte Logik aufgeklärt? Bitte fügen Sie Ihrer Antwort die entsprechenden Texte der Aufklärung bei.
- d. Auf welche Weise können die Informationen zur Datenverarbeitung bei der Nutzung des Dienstes aufgerufen werden?
- e. Ist nach der Nutzung von ChatGPT eine Datenlöschung nach Art. 17 DSGVO möglich?
- f. Gibt es automatisierte Entscheidungen (einschließlich Profiling) nach Art. 22 DSGVO durch OpenAI oder andere Unternehmen? Wenn ja, wie wird Art. 22 DSGVO umgesetzt?

IV. Datensicherheit, Data Protection by Design, Data Protection by Default

- a. Welche technischen und organisatorischen Maßnahmen werden bei ChatGPT eingesetzt, um die Risiken für die Rechte und Freiheiten der Benutzenden einzudämmen?
- b. Werden personenbezogene Daten pseudonymisiert, anonymisiert oder auf andere Weise aufbereitet, bevor diese zum Training der KI verwendet werden? Bitte beschreiben Sie das Verfahren.
- c. Wird bei der Erhebung und Verarbeitung unterschieden, ob die genutzten Daten einen Personenbezug aufweisen oder nicht? Wie wird das Risiko der Re-Identifizierbarkeit infolge des Trainings der KI und eventuell erstellter Profile bewertet?
- d. Besteht ein Sicherheitskonzept, mit dem nachgewiesen werden kann, dass solche Maßnahmen getroffen wurden? Wenn ja, bitte legen Sie das Dokument vor.

- e. Wie wird sichergestellt, dass bei der Nutzung des Dienstes erhobene personenbezogene Daten der Nutzenden (z.B. Namen, Telefonnummern, E-Mail-Adressen, Zahlungsinformationen, Nutzendeneinhalte/„usage data“ nach der Privacy Policy von OpenAI) nicht im Ergebnistext anderer Nutzenden genannt werden? Werden hierzu Anonymisierungs- oder Pseudonymisierungsverfahren angewandt? Bitte fügen Sie Ihrer Antwort die entsprechenden Verfahrensbeschreibungen bei.
- f. Wie werden bei ChatGPT und den von ChatGPT genutzten Generatoren GPT-3 und GPT-4 die Grundsätze des „Data Protection by Design“ und „Data Protection by Default“ umgesetzt? Bitte fügen Sie Ihrer Antwort Ihr Datenschutzkonzept mit den entsprechenden Erwägungen bei.

V. Datenschutz-Folgenabschätzung (DS-FA)

- a. Wurde eine DS-FA nach Art. 35 DS-GVO durchgeführt? Wenn ja, bitte legen Sie den DS-FA-Bericht vor.
- b. Zu welchem Zeitpunkt wurde die DS-FA durchgeführt?
- c. Sollte keine DS-FA durchgeführt worden sein, bitte ich um Erläuterung der für diese Entscheidung zugrundeliegenden Risikoabschätzung.

VI. Besonderer Schutz von Kindern und Jugendlichen:

- a. Welche Altersgrenze gilt für die Nutzung von ChatGPT?
- b. Wie stellen Sie sicher, dass der Dienst nicht von Personen unterhalb dieser Altersgrenze genutzt wird?
- c. Auf welche Weise werden ggf. erforderliche Einwilligungen von Minderjährigen eingeholt?

VII. Übermittlung von Daten außerhalb der Europäischen Union:

- a. Wo befinden sich die Server für den Dienst ChatGPT?
- b. Wo werden die Daten der betroffenen Personen aus Schleswig-Holstein physisch verarbeitet?
- c. Werden von Ihnen personenbezogene Daten von Nutzenden aus Schleswig-Holstein an andere Unternehmen mit Sitz außerhalb der Europäischen Union übermittelt?
- d. Welche von den in Frage VII c. benannten Unternehmen sind von Ihnen beauftragte Unterauftragnehmer/innen und welche Unternehmen sind verbundene Unternehmen („affiliates“) nach Nr. 3 der Privacy Policy von OpenAI?

VIII. Nutzung durch andere Dienste oder Unternehmen

- a. In welchen anderen Produkten oder Diensten anderer Unternehmen ist der Dienst ChatGPT bzw. die diesem zugrundeliegende KI eingebunden?

- b. Wird bei diesen Diensten jeweils der gleiche Funktionsumfang angeboten bzw. beruhen die Dienste auf dem gleichen Umfang der Datenverarbeitung?
- c. Werden im Rahmen der unter a) genannten Dienste personenbezogene Daten von Dritten erhoben, die von Ihrem Unternehmen zu eigenen Zwecken genutzt werden (z.B. zur Analyse, Training der KI)? Falls ja, auf welche Rechtsgrundlage stützen Sie eine solche Verarbeitung personenbezogener Daten?
- d. Wer ist bei der Nutzung durch andere Dienste oder Unternehmen Verantwortlicher nach Art. 4 Nr. 7 DS-GVO? Erfolgt die Nutzung durch Dritte im Rahmen einer Auftragsverarbeitung nach Art. 28 DS-GVO oder im Rahmen einer gemeinsamen Verantwortlichkeit nach Art. 26 DS-GVO? Bitte legen Sie entsprechende Musterverträge vor.

Bitte lassen Sie mir Ihre Antwort bis zum **7. Juni 2023** zukommen.

Anmerkung vom 01.06.2023:
Es wurde Fristverlängerung
beantragt und auch gewährt.

Ich weise Sie gemäß § 40 Abs. 4 S. 2 BDSG darauf hin, dass Sie die Auskunft auf solche Fragen verweigern können, deren Beantwortung Sie selbst oder eine/einen der in § 383 Abs. 1 Nr. 1 bis 3 der Zivilprozessordnung (ZPO) bezeichnete/n Angehörige/n der Gefahr strafgerichtlicher Verfolgung oder eines Verfahrens nach dem Gesetz über Ordnungswidrigkeiten (OWiG) aussetzen würde. Sollten Sie von dem Auskunftsverweigerungsrecht Gebrauch machen wollen, sind Sie verpflichtet, mir dies mitzuteilen.

Mit freundlichen Grüßen

Dr. h. c. Marit Hansen
Landesbeauftragte für Datenschutz Schleswig-Holstein