KI:connect.NRW

Fragen & Antworten

zur Nutzung des Microsoft Azure OpenAI Service
in der Version 1.0 von 2024-07-02

Kommentiert [MP1]: Das Dokument enthält stellvertretende Einträge, z.B. das Logo im Dokumentenkopf oder Datumsangaben. Alle Stellen mit solchen Einträgen sind über Kommentare mit dem Text „Anpassen.“ markiert.

Auch in den übrigen Passagen kann sich Änderungsbedarf für die eigene Hochschule ergeben. Das Dokument sollte dann entsprechend angepasst werden.

Generative KI
an Hochschulen in NRW

Executive Summary

1) **Art der verarbeiteten Daten:** Bei der Nutzung von GPT werden (i) die explizite Chat-Kommunikation zwischen Nutzenden und dem KI-Modell, (ii) personenbezogene Daten, also (Teil-)Informationen auch in pseudonymisierter Form, die zur Identifizierung einer Person dienen, sowie (iii) nicht-personenbezogene Daten verarbeitet.
2) **Verantwortliche für die Datenverarbeitung:** Die Verarbeitung all dieser Daten findet immer und ausschließlich innerhalb der IT-Infrastrukturen der Hochschule und/oder von Microsoft statt.
3) **Auftragsdatenverarbeiterin:** Verarbeitet werden die Daten durch die in der EU ansässige Microsoft Irland Operations Limited. Insbesondere werden keine Daten an OpenAI, Inc. und/oder OpenAI, L.P., übertragen.
4) **Verbindung zu OpenAI:** Microsoft betreibt eigene Instanzen von GPT vollständig auf dedizierten Servern von Microsoft. Diese werden im Rahmen der „Microsoft Azure OpenAI Services“ bereitgestellt. Zu keinem Zeitpunkt werden Daten aus der Hochschule durch Microsoft an OpenAI oder auch an andere Datenverarbeiterinnen bereitgestellt. Insbesondere werden auch in keinem Fall Inhalte von Eingaben („Prompts“) für Trainingszwecke verwendet.
5) **Datenverarbeitung in der EU:** Die Hochschule hat für alle an der Hochschule eingesetzten Azure Services und damit auch für das KI-Modell die Verarbeitung und Speicherung der Daten ausschließlich innerhalb der EU gewählt. Dies entspricht den Regelungen der EU Data Boundary for the Microsoft Cloud, kurz EUDB.
6) **Personenbezogene Daten:** Durch die Nutzung einer Hochschul-eigenen Weboberfläche und die Anmeldung im Rahmen des Single Sign-On werden keinerlei personenbezogene Daten an Dritte übertragen, solange die Nutzungsbedingungen der Hochschule eingehalten werden.
7) **Löschung übertragener Daten:** Alle übertragenen nicht-anonymen Daten werden auf Seiten von Microsoft nur vorübergehend gespeichert und mit dem Ende der Verarbeitung einer Eingabe („Prompt“) gelöscht. Ausnahmen sind hier die Eingabe selbst sowie vom KI-Modell generierte Ausgaben. Diese werden gemäß der „Responsible AI“ Prinzipien seitens Microsoft für 30 Tage gespeichert, um Missbrauch zu erkennen und nachträglich verfolgen zu können. Daten, die zum Betrieb durch die Hochschule aktiv gespeichert werden (z.B. Accountdaten, Logdaten) werden nach einem festgelegten Zeitraum gelöscht.
8) **Verbotener Content:** Gemäß des Code of Conduct for Microsoft Azure OpenAI Services werden potenziell schädliche Inhalte wie etwa Hatespeech oder gesundheitsgefährdende Anfragen regulatorisch untersagt und technisch aktiv unterbunden. Die Hochschule kann Ausnahmen selbst flexibel festlegen. Es findet aber keine automatische Sanktionierung der Verursachenden statt.

Kommentiert [MP2]: Anpassen.

Vorbemerkung

Im Folgenden wird allgemein von "Daten" und ihrer Verarbeitung gesprochen. Grundsätzlich sind damit drei verschiedene Arten von Daten gemeint:

D1. **Eingaben, Ausgaben und Aktionen:** Dies sind explizite Kommunikationsdaten, die zwischen Nutzenden und dem KI-Modell ausgetauscht werden. Dem unterfallen sowohl die Eingaben (Input) durch Nutzende im Rahmen eines Chats wie z.B. ein Prompt und die Ausgaben des KI-Modells (Output) als auch Aktionen wie das Löschen oder Umbenennen eines Chats.

D2. **Personenbezogene Daten:** dies sind Informationen oder Teilinformationen auch in pseudonymisierter Form, die zur Identifizierung einer Person dienen.

D3. **Nicht-personenbezogene Daten:** dies sind Informationen oder Teilinformationen, die nicht unter D1 oder D2 fallen, z.B. Metadaten oder anonymisierte Logdaten.

Fragenkatalog

Nachstehend wird ein Katalog von Fragen formuliert und beantwortet, die für den Betrieb der Azure OpenAI Services an der Hochschule maßgeblich sind. Die Beantwortung der Fragen ist im gemeinsamen Austausch zwischen der Hochschule und Microsoft erfolgt.

F1. Welche Daten werden verarbeitet und bei welchen Anlässen geschieht dies:

Generell ist die Art der verarbeiteten Daten abhängig von der gewählten Authentifizierungsart. Bei Multifaktor-Authentifizierung wird beispielsweise eine Telefonnummer für die SMS-Bestätigung verarbeitet. Die Authentifizierung der an der Hochschule genutzten Azure OpenAI Services wird nach aktuellem Planungsstand immer und ausschließlich über die Hochschule Tenants stattfinden (M365 und Azure). Informationen zum Konzept des Tenant (zu deutsch: „Mandant“) liefert: https://learn.microsoft.com/en-us/microsoft-365/solutions/tenant-management-overview.

Die nachstehenden Informationen beziehen sich auf die Authentifizierungsvariante über Hochschule Tenants.

a. Anlass: beim Login über Hochschule Tenants,

Durch die Nutzung einer von der Hochschule on-premise bereitgestellte Weboberfläche und des Hochschule Single Sign-On werden keinerlei personenbezogene Daten mehr an Microsoft übertragen.

b. Anlass: beim Absenden bzw. Bearbeiten eines Prompts,

Neben dem Payload der Anfrage (z.B. der Text des Prompts) werden betriebsnotwendige nicht-personenbezogene Metadaten übertragen, z.B. eine

anonyme *Chat ID* zur Identifikation eines Chats oder eine *Prompt ID* zur Identifikation eines Prompts.

c. Anlass: beim Auslösen von Begleitfunktionen, z.B. Thumbs-Up/Down oder Setzen von Einstellungen?

Neben dem Payload der Anfrage (z.B. Aktionstrigger wie „Neuer Chat“, „Daumen hoch“ oder „Chat löschen“) werden betriebsnotwendige nicht-personenbezogene Metadaten übertragen, z.B. eine anonyme *ChatID* zur Identifikation eines Chats oder eine *PromptID* zur Identifikation eines Prompts.

F2. Welche dieser Daten werden nicht-flüchtig gespeichert?

Unter einer nicht-flüchtigen Speicherung verstehen wir eine Speicherung über den unmittelbaren Verarbeitungsvorgang der gesendeten Daten hinaus. Sobald also die mit den primären Verarbeitungszwecken einer Anfrage verbundenen IT-Prozesse beendet sind, beginnt die nicht-flüchtige Speicherung. Die Konzeption von GPT ist explizit auf eine Minimierung der nicht-flüchtigen Speicherung ausgelegt.

Kommentiert [MP3]: Anpassen.

Bei nicht-flüchtigen Daten ist generell zu unterscheiden zwischen

(A) einer Speicherung auf Servern von Microsoft,

(B) einer Speicherung innerhalb der Hochschule Tenants, die aber auf Cloud-Servern erfolgt,

(C) einer Speicherung innerhalb der Hochschule auf Servern an der Hochschule (sog. „on-premise“ Variante).

Eine Speicherung personenbezogener Daten nur noch in Form von (C) statt. Es werden keinerlei personenbezogene Daten an Dritte übertragen, solange die Nutzungsbedingungen der Hochschule eingehalten werden.

Für nicht-personenbezogene Daten im Rahmen der Eingaben durch die Nutzenden findet gemäß der „Responsible AI“ Prinzipien seitens Microsoft eine anonymisierte Speicherung (i) der Prompts an das KI-Modell sowie (ii) der Ausgaben durch das KI-Modell in Form von (A) für 30 Tage statt, um Missbrauch zu erkennen und nachträglich nachgehen zu können.

Weitere Daten im Sinne von (A), die außerhalb der Hochschule Tenants gespeichert werden, sind ausschließlich anonymisierte Metadaten über die Nutzung des Dienstes. Dies sind Diagnosedaten wie zum Beispiel das benutzte KI-Modell, die Anzahl der übertragenen Tokens/Zeichen oder an welchem Datum die Übertragung erfolgt ist. Die Anonymisierung findet dabei oberhalb der Tenant-Ebene statt, so dass die gespeicherten Daten bereits nicht auf die Hochschule als Institution zurückzuführen sind und keinesfalls auf einzelne Nutzende rückschließen lassen.

F3. Welche konkreten Verarbeitungs- bzw. Speicherorte gibt es außerhalb der Hochschule? Insbesondere

a. innerhalb der IT-Infrastruktur von Microsoft in der EU,

Grundsätzlich findet die Verarbeitung von Daten, die nicht innerhalb der Hochschule Tenants verarbeitet werden (vgl. F2), immer und vollständig innerhalb der IT-Infrastruktur von Microsoft statt. Insbesondere werden keine Daten an OpenAI, Inc. und/oder OpenAI, L.P., übertragen. Wird eine Bereitstellung innerhalb der EU gewählt, dann findet die Verarbeitung und Speicherung in der Microsoft IT-Infrastruktur innerhalb der EU statt (EUDB). Mehr Informationen liefert F4.

b. innerhalb der IT-Infrastruktur von Microsoft außerhalb der EU,

Die Bereitstellung innerhalb der EU gemäß F4.a gilt nur eingeschränkt bei Nutzung des geplanten Services durch Personen, die sich von außerhalb der EU anmelden. Auch dort findet die Verarbeitung aber immer und vollständig innerhalb der IT-Infrastruktur von Microsoft statt.

c. außerhalb von a. oder b.?

Eine Verarbeitung außerhalb der Microsoft IT-Infrastruktur findet nicht statt. Microsoft betreibt eine komplett eigenständige „Kopie“ von ChatGPT vollständig innerhalb der eigenen Serverinfrastruktur. Das ist Teil der Unternehmenskooperation zwischen OpenAI Ltd. und Microsoft. Insbesondere nutzt Microsoft für die Azure OpenAI Services in keinem Fall die Server von OpenAI, sondern betreibt eine eigene GPT-Instanz. Es fließen bei der Nutzung der Azure OpenAI Services entsprechend keinerlei Daten aus der Hochschule bzw. von Microsoft an Dritte ab.

F4. Ort und Betriebsverantwortung für die Verarbeitung und Speicherung dieser Daten:

a. Für welche der Daten aus F1 kann gewährleistet werden, dass Verarbeitung und Speicherung in Datacentern innerhalb der EU ablaufen?

Die Hochschule hat für die Bereitstellung der Ressourcen für alle Azure Services mindestens den Serverstandort „Europa“ festgelegt, so dass keine Datenverarbeitung auf Servern außerhalb der EU erfolgt. Der Serverstandort „Europa“ bezeichnet dabei zunächst die von Microsoft so genannte „Makroregion Geografie 1 – EMEA“. Diese umfasst Rechenzentren in Ländern wie Österreich, Finnland, Frankreich, Irland, Niederlande, Polen und Schweden. Zusätzlich kann die Hochschule für bestimmte Services den Standort „Geografie der lokalen Region“ wählen, womit üblicherweise eine Bereitstellung durch Server im jeweiligen Land gemeint ist. Für die Bundesrepublik Deutschland sind dies dann Rechenzentren im Bundesgebiet. Mehr Informationen liefert: https://learn.microsoft.com/de-de/microsoft-365/enterprise/m365-dr-overview?view=o365-worldwide

Bei dieser Datenspeicherung gibt es Ausnahmen, die sich vor allem auf den Authentifizierungsvorgang beziehen. Da die Hochschule eine Authentifizierung der Nutzenden rein über die so genannte „Microsoft Entra ID“ (früher Azure Active Directory) durch die Nutzung des Single Sign-On Verfahrens technisch ausschließt, beziehen sich die Ausnahmen ausschließlich auf eine Nutzung der Multifactor

Authentifizierung (MFA) sowie auf den Zugriff von einem Ort *außerhalb* der Microsoft European Data Boundary (EUDB). Die Ausnahmen sind hier weiter dokumentiert: https://www.microsoft.com/de-de/trust-center/privacy/european-data-boundary-eudb

Die Hochschule empfiehlt überdies in ihren Nutzungsbedingungen, auf die an der Hochschule betriebenen Azure OpenAI Services nur von innerhalb der Hochschule bzw. aus dem Hochschule VPN zuzugreifen, wodurch diese vorgenannten Ausnahmen weiter minimiert werden.

b. Wer ist Betreiber der Datacenter in der EU gemäß F4.a?

Betreiber ist in jedem Fall Microsoft Irland Operations Limited.

c. Für welche Daten ist die Antwort zu F4.a nichtzutreffend?

Es gibt keine Daten, die hierunter fallen. Microsoft Irland Operations Limited ist Betreiberin der Datacenter, unabhängig von den Daten.

F5. **Werden die Azure OpenAI Services vollständig in Datacentern bzw. auf Servern von Microsoft ausgeführt oder werden die für den Betrieb der generativen KI erforderlichen Services ganz oder teilweise auf Servern der Firma OpenAI bzw. auf anderen Servern ausgeführt**?

Es wird keine Infrastruktur von OpenAI (OpenAI, Inc.' und/oder OpenAI, L.P.) verwendet. Alle Verarbeitung geschieht auf Servern von Microsoft (vgl. die Antwort zu F4.a und F4.c).

Dies kann sich ändern, sofern durch die Hochschule andere Services hinzugezogen/eingebunden werden. Von Seiten der Hochschule ist das aktuell nicht der Fall, nicht geplant und würde die üblichen Mitbestimmungsprozesse erfordern.

F6. Werden Prompts der Nutzenden, die im Rahmen der Nutzung der Azure OpenAI Services entstehen, zu Trainingszwecken bei der Weiterentwicklung des Sprachmodells durch Microsoft oder Dritte genutzt? Um welche Dritten handelt es sich dabei?

Es findet durch Microsoft oder Dritte in keinem Fall eine Nachnutzung der Eingaben der Nutzenden zu Trainingszwecken statt.

F7. Was geschieht im Falle der Löschung eines Chatverlaufs bzw. aller Chatverläufe mit den im Rahmen des Chatverlaufs übertragenen Daten?

Generell sind Chats persistent, d.h. sie sind auch nach Tagen oder Wochen für die Nutzenden noch einsehbar. Dies dient der Nachvollziehbarkeit und Wieder-Aufrufbarkeit von Chatverläufen durch Nutzende und ist eine Kernfunktionalität des KI-Dienstes. Im Rahmen des Pilotbetriebs soll geklärt werden, welche „Verfallszeit" für Chats einstellbar sein soll, so dass Chats nach einer vorgegebenen Zeitdauer wieder gelöscht werden.

Die Löschung eines Chats bzw. aller Chats erfolgt über den Klick auf den jeweiligen Button in der Weboberfläche. Wird ein Löschen-Button geklickt, wird abhängig vom Button der aktuelle Chat oder alle Chats inklusive damit verbundener Daten auf den Microsoft-Servern gelöscht und kann nicht wiederhergestellt werden.

F8. Wie werden Löschanfragen von Nutzenden umgesetzt?

Löschanfragen müssen Nutzende an den Data Controller stellen: Das ist im vorliegenden Fall die Hochschule. Der Data Controller nutzt bei der Umsetzung von Löschanfragen die im Rahmen der übergreifenden Azure-Funktionalität bereitgestellten Werkzeuge. Microsoft garantiert die technische Umsetzung der Löschanfrage gemäß der im Service Agreement beschriebenen Bedingungen.

F9. Wie werden Auskunftsersuchen von Nutzenden umgesetzt?

Auskunftsersuchen müssen Nutzende an den Data Controller stellen. Das ist im vorliegenden Fall die Hochschule. Der Data Controller nutzt bei der Umsetzung von Auskunftsersuchen die im Rahmen der übergreifenden Azure-Funktionalität bereitgestellten Werkzeuge. Microsoft garantiert die technische Umsetzung der Auskunftsersuchen gemäß der im Service Agreement beschriebenen Bedingungen.

F10. Wie ist die Übertragung der Nutzungsrechte an Chatverläufen geregelt?

Auf Seiten von Microsoft besteht kein Eigentum an Kundendaten. Dies ist im bestehenden Data Processing Agreement im Rahmen der Hochschule Azure Services beschrieben. Dazu gehören alle oben beschriebenen Datenarten D1, D2 und D3.

F11. Bei der Nutzung von https://chat.openai.com/ existieren Terms and Conditions sowie damit verbundene technische und organisatorische Maßnahmen, die bestimmte Nutzungszwecke unterbinden. So werden z.B. das Generieren von "Hatespeech" oder einer medizinischen Beratung sowohl regulatorisch untersagt als auch technisch aktiv unterbunden, sobald die KI einen solchen Zweck erkennt. Welche dieser Restriktionen sind "out-of-the-box" auch bei den Azure OpenAI Services implementiert?

Auch in den Azure OpenAI Diensten werden Hatespeech etc. per Default regulatorisch untersagt und technisch aktiv unterbunden. Überdies lassen sich Ausnahmen selbst flexibel in Azure festlegen. Im Falle von Verstößen findet nach derzeitigem Stand keine regelmäßige, automatische und dauerhafte Sanktionierung der verursachenden Person statt, z.B. eine zeitlich unbegrenzte Sperrung des Zugangs zum KI-Modell.

Hier ein Überblick und Startpunkt: https://learn.microsoft.com/en-us/legal/cognitive-services/openai/code-of-conduct