Hochschule für
Wirtschaft und Recht Berlin
Berlin School of Economics and Law

Merkblatt des Datenschutzbeauftragten

Merkblatt zur Nutzung von Generativer KI und Datenschutz

(Stand 04/2024)

In diesem Merkblatt wird aufgezeigt, welche Daten durch die Nutzung von generativer KI erhoben und verarbeitet werden und wie man generative KI datenschutzkonform nutzen kann. Exemplarisch soll hier insbesondere der Chatbot ChatGPT der Firma OpenAI näher beleuchtet werden.

Das Merkblatt kann jedoch **keine eingehende rechtliche Prüfung des Einzelfalles ersetzen**. Im Zweifelsfall können Sie sich mit konkreten Fragen gern an den Datenschutzbeauftragten wenden. Sorgfalt und Transparenz in Sachen des Datenschutzes sind in jedem Fall geboten und zahlen sich langfristig aus, da Rechtsunsicherheiten und Rechtsverstöße vermieden werden. Auch im Rahmen der Nutzung von generativer KI müssen die datenschutzrechtlichen Bestimmungen der Datenschutz-Grundverordnung (DGSVO) sowie ergänzender Datenschutznormen beachtet werden.

Zwei zentrale Empfehlungen sollen vorweg gegeben werden:

- *Geben Sie nur wenige persönliche Daten preis, u.a. weder private Accounts und E-Mail-Adressen noch ihr korrektes Geburtsdatum.*
- *Vertrauliche Informationen sollten nicht über ChatGPT geteilt werden.*
- *Geben Sie keine urheberrechtlich geschützten Informationen in die generative KI ein.*

1. Welche Daten von Nutzenden werden bei Verwendung generativer KI verarbeitet?

Welche Daten verarbeitet werden, unterscheidet sich zum Teil stark von KI-Dienst zu KI-Dienst und hängt unter Umständen auch davon ab, ob man die kostenfreie Basis- oder die Bezahlvariante des jeweiligen Dienstes nutzt.

1.1. Registrierungsdaten

Im Rahmen des Registrierungsprozederes werden sowohl für die Bezahl- als auch die Basisversion E-Mail-Adresse sowie die Mobilnummer der Userinnen und User abgefragt. Daneben werden bei der Bezahlversion auch Bezahldaten abgefragt.

1.2. Nutzungsdaten

In den Logfiles werden die üblichen Daten protokolliert, u.a. IP-Adressen und Geräteinformationen, vermutlich auch Accountnamen und ausgeführte Aktionen. Wie lange die Daten gespeichert werden und was sonst mit diesen im Nachgang geschieht, ergibt sich bislang aus der Datenschutzerklärung von OpenAI nicht.

1.3. Chatinhalte

Die Inhalte der jeweiligen Eingabe, also die Prompts, werden ebenfalls gespeichert und verarbeitet. Dies können Texte, Bilder und sonstige Daten und Eingaben sein.

Eine Möglichkeit, datensparsamer gegenüber individuellen Registrierungen der Lehrkräfte und Studierenden zu agieren, besteht darin, Schnittstellen-Lösungen in den digitalen Lernumgebungen der Hochschulen zu integrieren. Diese ermöglichen eine pseudonyme Nutzung und können je nach

Ausgestaltung auch regelmäßige Löschungen der bei Nutzung der Dienste verarbeiteten Daten vorsehen. Ein aktuelles Beispiel dafür ist HAWKI. Dabei handelt es sich um ein didaktisches Interface für Hochschulen, das auf der API von OpenAI basiert. An der HWR Berlin wird dieser Dienst bislang noch nicht angeboten. Der Hinweis darauf dient dafür, aufzuzeigen welche Möglichkeiten es insgesamt gibt.

Voreingestellt ist die dauerhafte Speicherung der Chathistorie, die von OpenAI zum Trainieren des Large Language Models (LLM) genutzt wird. In welcher Form die Chats dabei übernommen werden, ist nicht offensichtlich. Zudem können bereits eingeflossene Inhalte in die KI von OpenAI bislang nicht per Widerruf gelöscht werden.

Die Voreinstellung kann jedoch in den Einstellungen von OpenAI („Chat history & training“ unter „Data Controls“) deaktiviert werden, sodass die Inhalte den Nutzenden nicht mehr angezeigt und nach 30 Tagen von OpenAI gelöscht werden.

2. Auf welche Rechtsgrundlage lässt sich der Einsatz von KI-Diensten wie ChatGPT in der Hochschullehre stützen?

Der Einsatz von KI-Anwendungen in der Lehre darf grundsätzlich nur unter Einhaltung geltender datenschutzrechtlicher Bestimmungen erfolgen. Es sind insbesondere die in Art. 5 DSGVO festgelegten Grundsätze für die Verarbeitung personenbezogener Daten zu beachten. Darunter fällt auch, dass die Nutzung, also die Verarbeitung personenbezogener Daten, rechtmäßig erfolgen muss. Das bedeutet, dass eine Rechtsgrundlage vorliegen muss, die die Datenverarbeitung an der Hochschule erlaubt (vgl. Art. 5 Abs. 1 lit. a DSGVO).

An Hochschulen kommt derzeit allenfalls eine freiwillige Nutzung auf Basis einer informierten Einwilligung gemäß Art. 6 Abs. 1 S. 1 lit. a DSGVO in Betracht. Das heißt bedeutet, dass Studierende im Rahmen von Lehrveranstaltungen von Lehrenden nicht zur Nutzung, insbesondere nicht zur Registrierung, verpflichtet werden dürfen, da sonst die Freiwilligkeit nicht mehr vorliegt. Eine informierte Einwilligung setzt weiterhin voraus, dass die jeweiligen Studierenden die Gelegenheit haben, transparente Datenschutzhinweise zur Kenntnis zu nehmen. Zudem muss den Studierenden eine faktische Alternative angeboten werden, z. B. die Nutzung eines alternativen datenschutzkonformen KI-Dienstes aus Deutschland oder der EU oder die Teilnahme an einem gleichwertigen Seminar.

Auch zu empfehlen ist der Rückgriff auf Schnittstellen bzw. Plugins, die zumindest eine pseudonyme Nutzung ermöglichen, so dass personenbezogene Daten nur eingeschränkt verarbeitet werden oder die Nutzung über den Zugang der Lehrperson.

3. Welche datenschutzrechtlichen Verpflichtungen kommen auf Hochschulen bei Nutzung generativer KI-Modelle zu?

Hinsichtlich des Datenschutzes bei Einsatz von KI-Diensten bestehen die üblichen Handlungsbedarfe wie bei sonstigen digitalen Diensten und Anwendungen, die an Hochschulen zu Lehrzwecken eingesetzt werden und personenbezogene Daten verarbeiten.

Zu den wichtigsten Maßnahmen zählen:

- die Überprüfung des Vorliegens einer Rechtsgrundlage, die die Nutzung zu Lehrzwecken gestattet;
- die Vornahme einer allgemeinen Datenschutzprüfung inkl. Prüfung der DSGVO-Konformität der technischen und organisatorischen Maßnahmen (sog. TOMs), die die Sicherheit der

Verarbeitung personenbezogener Daten gewährleisten sollen (Bsp.: Sicherstellung von Schutzmaßnahmen gegen den Zugriff Dritter);
- ggf. der Abschluss eines Auftragsverarbeitungsvertrages (oder, je nach Konstellation, einer Vereinbarung über die gemeinsame Verantwortlichkeit)
- das Verfassen und Bereitstellen von Datenschutzhinweisen für die Nutzung;
- die Dokumentation aller oben genannten Schritte und Vertragsunterlagen im sog. Verzeichnis von Verarbeitungstätigkeiten.

Mit Hinblick auf mögliche Risiken für die Rechte und Freiheiten betroffener Personen durch Datenverarbeitungen im Zuge der KI-Nutzung in der Lehre wird in der Regel auch eine sog. Datenschutzfolgeabschätzung i. S. v. Art. 35 DSGVO vorzunehmen sein. Es handelt sich dabei um eine Risikobewertung, innerhalb derer geprüft und eingeschätzt wird, welche Folgen die Verarbeitung personenbezogener Daten im Zuge der Nutzung des gewählten KI-Dienstes für den Schutz der Daten der hiervon Betroffenen haben kann.

4. Fazit

Der Einsatz KI-basierter Text- und Bildgeneratoren an Hochschulen bzw. im Rahmen der Hochschullehre kann unter Einhaltung datenschutzrechtlicher Bestimmungen durchaus möglich sein. Bislang sind jedoch nicht alle Fragen in diesem Rahmen geklärt, sodass nach wie vor Vorsicht geboten ist. Alle Nutzerinnen und Nutzer sollten die Dienste bewusst und verantwortungsvoll nutzen, fremde Rechte dabei stets beachten und KI-Output vor Weiternutzung auch immer kritisch hinterfragen.

Im Zweifel sollte immer der Datenschutzbeauftragte kontaktiert werden, um etwaige Datenschutzverstöße zu vermeiden.