Deep Dive

Datenschutz und KI

Risiken erkennen, Chancen nutzen

Niklas Mühleis ist Rechtsanwalt und Partner der Kanzlei Heidrich Rechtsanwälte in Hannover Joerg Heidrich ist Justiziar und Datenschutzbeauftragter von Heise Medien, Fachanwalt für IT-Recht und Partner der Kanzlei Heidrich Rechtsanwälte

Inhalt

1. KI und Datenschutz: Eine Einführung
   1.1. KI und Datenschutz: Geht das überhaupt?
   1.2. Personenbezogene Daten: Was KI-Nutzer wissen müssen
   1.3. Personenbezogene Daten im KI-Kontext
   1.4. Vorsicht mit sensiblen Daten
   1.5. Anonymisierung und Pseudonymisierung von Daten
2. Verantwortlichkeiten und Rollenverteilung
   2.1. Wer ist verantwortlich? Rollenverteilung zwischen Unternehmen und KI-Anbietern
3. Rechtsgrundlagen für den KI-Einsatz
   3.1. Rechtsgrundlagen im Überblick

Veröffentlicht: 12. Mai 2025

KI und Datenschutz: Eine Einführung

KI und Datenschutz: Geht das überhaupt?

Künstliche Intelligenz (KI) verändert Wertschöpfungsketten in rasantem Tempo. Unternehmen profitieren von Kostenreduktion, höherer Geschwindigkeit und neuen Geschäftsmodellen. Ob automatisch verfasste Produktbeschreibungen, intelligente Chatbots im Kundenservice oder Code-Vervollständigung für Entwicklerteams: Die Einsatzmöglichkeiten scheinen nahezu grenzenlos und versprechen Effizienzgewinne.

Doch so groß die Faszination ist, so laut klingeln bei internen Datenschützern und den staatlichen Aufsichtsstellen die Alarmglocken. Denn spätestens so bald ein solches System mit persönlichen Daten gefüttert wird – seien es Kundendaten oder Mitarbeiterinformationen– tritt die strenge Datenschutz-Grundverordnung (DSGVO) auf den Plan.

Und diese verträgt sich bereits im Grundsatz nicht besonders gut mit generativer KI. Denn der Datenschutz sieht eine größtmögliche Information und Transparenz bei der

Verarbeitung von Daten vor. Solche Grundsätze lassen sich aber gerade bei LLM-Modellen nur sehr schwer durchsetzen, denn diese sind von Haus aus Blackboxes. Auch Betroffenengrundrechte wie Auskunft und Löschung sind mit KI-Angeboten nur sehr schwer umzusetzen.

Neben diesen grundsätzlichen Problemen stellen sich aber auch ganz viele ganz praktische Fragen: Darf ein Unternehmen überhaupt Texte von einer KI schreiben lassen, wenn darin Kundennamen oder andere persönliche Angaben auftauchen? Wer trägt die Verantwortung, wenn ein externer KI-Dienst genutzt wird? Und wie lässt sich verhindern, dass sensible Daten in die falschen Hände geraten?

Wichtig: Typischerweise gibt es an drei Stellen Berührungspunkte zwischen dem Datenschutz und der KI: Beim Training, bei der Eingabe und bei der Ausgabe. Diese drei Punkt müssen zwingend bei der KI-Strategie berücksichtigt werden. In unserem Deep Dive beschränken wir uns primär auf die Probleme bei der Nutzung von KI.

In diesem Deep Dive werfen wir einen ausführlichen Blick auf die datenschutzrechtlichen Fallstricke beim Einsatz generativer KI in Unternehmen – und zeigen Wege auf, wie sich die neuen Technologien möglichst DSGVO-konform nutzen lassen. Oder anders gefragt: Wie also lassen sich Innovation und Privatsphäre unter einen Hut bringen?

Personenbezogene Daten: Was KI-Nutzer wissen müssen

Zunächst ein grundlegender Blick auf den Kern des Problems: Was genau sind **personenbezogene Daten** und warum sind sie im Zusammenhang mit KI so wichtig? Die DSGVO definiert diese Angaben als alle Informationen, auf deren Basis eine Person identifizierbar ist. Das klingt abstrakt, umfasst in der Praxis aber alles von Namen, Adressen und Kundennummern über Standortdaten oder IP-Adressen bis hin zu Aussagen und Textinhalten, die Rückschlüsse auf eine Person zulassen. Sobald also ein generatives KI-System solche Informationen erfasst, verarbeitet oder ausspuckt, bewegt man sich im Geltungsbereich der DSGVO.

> Praxisbeispiel: Lässt ein Kundendienst-Mitarbeiter sich von einer KI wie ChatGPT beim Formulieren einer Antwort helfen und kopiert dafür die Anfrage des Kunden (inklusive dessen Name oder Vertragsdetails) in das KI-Tool, werden personenbezogene Daten an einen externen Dienst übermittelt und durch diesen verarbeitet. Ähnlich verhält es sich, wenn eine HR-Abteilung Bewerbungsunterlagen durch ein KI-System analysieren lässt – die Lebensläufe enthalten eindeutig persönliche Angaben. Auch die Daten, auf deren Basis sich Mitarbeiter in Unternehmen bei den KI-Diensten einloggen, unterfallen dem Datenschutz.

Auf der anderen Seite: Ohne personenbezogene Daten gibt es aus rechtlicher Sicht kein Datenschutzproblem. Wer eine KI völlig isoliert mit reinen Fantasiedaten oder anonymisierten Informationen füttert, braucht die DSGVO nicht zu fürchten. Doch solche Fälle sind im Unternehmensalltag selten, etwa bei der Verwendung reiner Maschinendaten. Meist arbeiten KI-Anwendungen gerade deshalb so nützlich, weil sie

echten Input aus der Firmenumgebung bekommen – und der ist nun einmal oft personenbeziehbar. Die Grenze zwischen anonym und personenbeziehbar verschwimmt leicht: Schon ein Kundenfeedback-Text kann personenbezogen sein, wenn beispielsweise der Verfasser indirekt identifizierbar ist, etwa durch Erwähnung spezifischer Umstände oder über den genutzten Usernamen. Oder die genutzten Maschinendaten enthalten Informationen zu dem zuständigen Maschinenführer.

Greift die KI im Unternehmen auf personenbezogene Daten zu oder lässt man sich diese generieren, etwa als potentielle Kundenliste, muss jedes Unternehmen die Vorgaben der DSGVO einhalten. Das Unternehmen übernimmt in diesem Moment in der Regel die Rolle des „**Verantwortlichen**“ im Sinne der DSGVO. Das bedeutet, dass die Firma dafür verantwortlich ist, dass die Datenverarbeitung durch die KI rechtskonform abläuft. Diese Verantwortung kann man nicht einfach auf den Hersteller der Software abschieben. Selbst wenn ein externer Dienst genutzt wird, bleibt das Unternehmen der Herr der Daten und muss Rechenschaft über deren Verwendung ablegen können.

Was bedeutet das konkret? Im nächsten Schritt schauen wir uns die wichtigsten datenschutzrechtlichen Pflichten und Problemfelder an – von der Frage nach der richtigen Rechtsgrundlage über Zweckbindung und Transparenz bis hin zu speziellen Anforderungen bei automatisierten Entscheidungen. Immer mit dabei: praxisnahe Beispiele und Tipps, wie man die Hürden meistern kann.

Personenbezogene Daten im KI-Kontext

Der Anwendungsbereich der DSGVO ist also immer dann eröffnet, wenn KI-Systeme personenbezogene Daten verarbeiten.

> • Beispiel Kundenservice: Ein Chatbot, der Kundenanfragen beantwortet, verarbeitet typischerweise Namen, Kontaktdaten oder Gesprächsverläufe einzelner Kunden. Diese Daten beziehen sich eindeutig auf Personen – die DSGVO ist voll anwendbar.
>
> • Beispiel interne Analyse: Nutzt eine Personalabteilung KI, um Mitarbeiterdaten wie Leistungskennzahlen oder E-Mail-Verkehr zu analysieren, werden Daten von identifizierbaren Mitarbeitern verarbeitet. Auch hier greift der Datenschutz.

Unternehmen sollten daher zunächst genau prüfen, wo überall personenbezogene Daten in ihre KI-Anwendungen einfließen. Falls möglich, kann eine frühzeitige Anonymisierung oder Pseudonymisierung von Daten den Datenschutzaufwand reduzieren. Das ist aber in vielen Fällen leichter gesagt als getan: Pseudonymisierte Daten (z.B. durch Codes statt Klarnamen) gelten weiterhin als personenbezogen, wenn eine Re-Identifizierung möglich ist.

> Praxis-Tipp: Machen Sie sich bei jedem KI-Projekt bewusst, welche Datenarten verarbeitet werden und ob ein Personenbezug vorliegt. Diese Bestandsaufnahme ist die Basis für alle weiteren Datenschutzmaßnahmen.