LOSCHELDER

Newsletter Datenschutzrecht

Januar 2025

Sehr geehrte Damen und Herren,

wir wünschen Ihnen für 2025 alles Gute und freuen uns, dass Sie uns auch in diesem Jahr folgen – Danke für Ihr Interesse an unserem Newsletter! Melden Sie sich immer gerne für die Vertiefung oder auch, wenn wir Sie bei konkreten Herausforderungen individuell begleiten können.

Ein bedeutendes Thema wird 2025 der Datenschutz in KI-Anwendungen sein. Kurz vor Start der ersten Pflichten aus der KI-Verordnung starten wir daher auch passend in das neue Jahr mit einem KI-Schwerpunkt und berichten zu einer ersten relevanten Pflicht aus der KI-Verordnung, der KI-Kompetenz, sowie zu neuen Entwicklungen in Sachen KI und Datenschutz.

Was Sie in Ihrem Unternehmen veranlassen müssen, um die KI-Kompetenzpflicht aus Art. 4 KI-Verordnung ab dem 2. Februar 2025 umzusetzen, ist auch Gegenstand unseres ersten Webinars im neuen Jahr. Dieses veranstalten wir gemeinsam mit Marcel Pesch von der academy4.ai, der im Webinar auch das Schulungsangebot der academy4.ai vorstellen wird – die rechtlichen Bausteine dort dürfen wir beisteuern. Melden Sie sich gerne an – wie gewohnt kostenfrei für unsere Mandanten und Behörden- wie Unternehmensvertreter:

KI-Kompetenz: Neue Schulungspflicht für Personal & Co. ab dem 2. Februar 2025!

Jeder, der KI im beruflichen Umfeld anbietet oder betreibt, muss ab Februar 2025 für eine ausreichende KI-Kompetenz sorgen, und zwar beim eigenen Personal und auch bei allen Personen, die in seinem Auftrag die KI betreiben oder nutzen. Dazu verpflichtet Artikel 4 KI-Verordnung alle Unternehmen, unabhängig von ihrer Größe. Was genau hinter dieser Pflicht steckt, wie sie umgesetzt werden kann und welche Risiken bei Nichtbeachtung drohen, besprechen wir in unserem Webinar am 30. Januar 2025. Wir beleuchten die KI-Kompetenz rechtlich und – dank unserer Zusammenarbeit mit Marcel Pesch von der academy4.ai – auch operativ und technisch. Wir werden Ihnen zum Abschluss des Webinars auch das Programm der academy4.ai vorstellen, welches für die Basisschulungen zur KI-Kompetenz beste Voraussetzungen schafft. Wir stehen selbstverständlich für Ihre Fragen zur Verfügung und freuen uns auf intensive Diskussionen mit Ihnen!

Donnerstag, den 30. Januar 2025 – 18:00 bis 19:00 Uhr
Ihre Referenten: Marcel Pesch (academy4.ai) und
Dr. Kristina Schreiber (Loschelder)

Alle Webinare bieten wir kostenfrei an und freuen uns über Ihre Anmeldung unter webinare@loschelder.de. Das Webinar findet über Teams statt, der Einladungslink wird rechtzeitig vor der Veranstaltung bereitgestellt.

Zum Jahresstart gibt es darüber hinaus auch unabhängig von KI weitere relevante Datenschutzthemen, die wir für Sie aufbereitet haben: Eine wortwörtlich „bewegende" Datenpanne wurde Ende des vergangenen Jahres bei VW bekannt. Wie wichtig das Zusammenspiel aus Datenschutz und Datensicherheit ist, wurde hier noch einmal ganz besonders deutlich. Wir berichten.

Auch die Aufsichtsbehörden hatten zum Ende des Jahres noch einmal ordentlich zu tun – zum Wohle des Datenschutzes klingelten im Dezember die Kassen: Welche Bußgelder gegen Netflix, Apple und Coolblue verhängt wurden, lesen Sie in diesem Newsletter im „Zu guter Letzt".

Inhalt

KI-Kompetenz: Neue Pflicht für Unternehmen ab dem 2. Februar 2025

Im August 2024 ist die europäische KI-Verordnung (Verordnung (EU) 2024/1689, kurz „KI-VO") in Kraft getreten. Ziel dieser Verordnung ist es vor allem, Risiken zu reduzieren, die von KI-Systemen ausgehen können. Um das umzusetzen, kommen zwangsläufig zahlreiche neue Verpflichtungen auf Unternehmen zu, welche KI-Systeme anbieten oder betreiben, was bei der beruflichen Nutzung schnell der Fall ist. Eine erste Pflicht, die für fast alle im professionellen Umfeld gilt, ist die KI-Kompetenz: Danach ist sicherzustellen, dass Ihr Personal über ausreichende KI-Kompetenz verfügt. Doch was bedeutet das genau und wie gestaltet sich die Umsetzung dieser Vorgabe?

Die ersten beiden Pflichten aus der KI-Verordnung gelten EU-weit ab dem 2. Februar 2025: Neben dem grundsätzlichen Verbot bestimmter Praktiken, werden ab 2. Februar 2025 (Art. 113 UAbs. 3 lit a) KI-VO) unzählige Unternehmen verpflichtet, KI-Kompetenz in ihrem Unternehmen sicherzustellen. Genauer: Art. 4 KI-VO verpflichtet Anbieter und Betreiber von KI-Systemen Maßnahmen zu ergreifen, um nach besten Kräften sicherzustellen, dass ihr Personal und andere Personen, die in ihrem Auftrag mit dem Betrieb und der Nutzung von KI Systemen befasst sind, über ein ausreichendes Maß an KI Kompetenz verfügen. Betroffene Unternehmen werden somit schon deutlich vor Geltungsbeginn der meisten weiteren KI-VO-Vorschriften verpflichtet, KI-Kompetenz unter ihren Mitarbeitenden zu vermitteln.

KI-Kompetenz ist *„die Fähigkeiten, die Kenntnisse und das Verständnis, die es Anbietern, Betreibern und Betroffenen unter Berücksichtigung ihrer jeweiligen Rechte und Pflichten im Rahmen dieser Verordnung ermöglichen, KI-Systeme sachkundig einzusetzen sowie sich der Chancen und Risiken von KI und möglicher Schäden, die sie verursachen kann, bewusst zu werden"* (Art. 3 Nr. 53 KI-VO). Die KI-Nutzer sollen befähigt werden, fundierte Entscheidungen über den Einsatz von KI-Systemen zu treffen (Erwägungsgrund 20 KI-VO).

Aber was genau bedeutet das? Mitarbeitenden muss vermittelt werden, wie KI-Systeme korrekt eingesetzt werden, welche Technik dahintersteht, wie Ergebnisse entstehen und welche rechtlichen Rahmenbedingungen für den konkreten Einsatz gelten. Im Zentrum

steht dabei die Befähigung der Mitarbeitenden, fundierte Entscheidungen über KI-Systeme zu treffen:

- Wann und wo können KI-Systeme sinnvoll eingesetzt werden?
- Mit welchen Daten wird das KI-System betrieben bzw. welche Datenbasis wird verwendet?
- Wie sind Outputs zu bewerten und zu nutzen?
- Welche Chancen und welche Risiken bestehen, wie kann ich die Risiken minimieren?

Dies zu erreichen, ist eine Daueraufgabe. Panik ist damit nicht geboten: Es müssen nicht am 2. Februar 2025 KI-Profis in jeder Ecke des Unternehmens sitzen.

Aber: Unternehmen sollten die Pflicht auch nicht ignorieren. KI-Kompetenz zu schaffen ist kein Selbstzweck. Die Fähigkeit sichert den gewinnbringenden Einsatz von KI im Unternehmen, auch und gerade zugunsten des Unternehmens selbst.

Um KI-Kompetenz sicherzustellen, sollten sich Unternehmen auf drei Säulen konzentrieren:

- Fortbildungen und **Schulungen** für die Mitarbeitenden, die KI einsetzen oder bald einsetzen werden
- Bereitstellung von Ansprechpartnern im Unternehmen und Zuteilung von Verantwortlichkeiten, z.B. durch Benennung eines sachkundigen **KI-Beauftragten**
- KI-Governance mit einem Überblick über die genutzten KI-Systeme, KI-Richtlinien für ihren Einsatz und eine Unternehmensstrategie zu Chancen, Risiken und Leitlinien beim KI-Einsatz

Verpflichtete der KI-Kompetenz

Verpflichtete des Art. 4 KI-VO sind Anbieter und Betreiber eines KI-Systems. Von der KI-Kompetenzpflicht werden somit unzählige Unternehmen betroffen sein, da ein Unternehmen schon dann als Betreiber eines KI-Systems anzusehen ist, wenn es KI-Systeme im eigenen Unternehmen für berufliche Zwecke einsetzt.

Anbieter sind dabei diejenigen, die ein KI-System entwickeln oder entwickeln lassen und es unter ihrem eigenen Namen oder ihrer

Handelsmarke in Verkehr bringen oder in Betrieb nehmen (Art. 3 Nr. 3 KI-VO).

Betreiber sind gem. Art. 3 Nr. 4 KI-VO diejenigen, die ein KI-System in eigener Verantwortung verwenden, es sei denn, das KI-System wird nur im Rahmen einer persönlichen und nicht beruflichen Tätigkeit verwendet. Betreiber von KI-Systemen sind somit alle Unternehmen, die KI-Systeme gezielt für berufliche Zwecke einsetzen bzw. einsetzen lassen, aber auch dann, wenn sie den Einsatz aktiv dulden. Die Pflicht zur Sicherstellung von KI-Kompetenz gilt dabei unabhängig von der Unternehmensgröße. Mittlere und kleine Unternehmen werden daher auch verpflichtet KI-Kompetenz herzustellen.

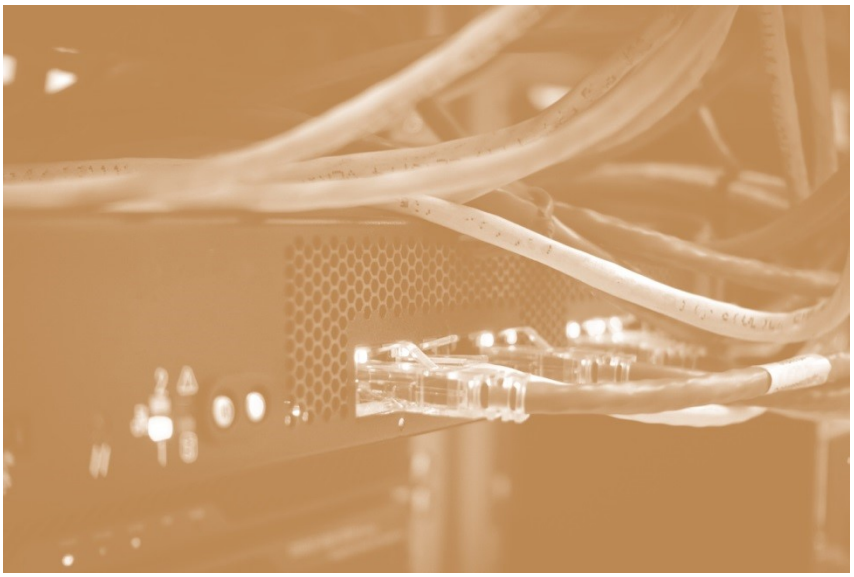
Umsetzung im Unternehmen

Für die Umsetzung und Sicherstellung der KI-Kompetenz gibt es - wie so oft - keine universelle Lösung. Der konkrete Umfang der jeweiligen KI-Kompetenz ist von verschiedenen Faktoren abhängig; unter anderem vom Umfang der Nutzung von KI-Systemen, Grad des Risikos des KI-Systems und der Position der Mitarbeitenden. Während alle Mitarbeitenden, die mit KI in Berührung kommen können, ein Grundverständnis von KI, dem Rechtsrahmen, technischen Gegebenheiten, Chancen und Risiken erhalten sollten, benötigen Verantwortliche für KI-Projekte zusätzlich fundierte Kenntnisse zu den umzusetzenden Pflichten der KI-VO.

Im ersten Schritt sollte daher Basiswissen hinsichtlich KI-Kompetenz vermittelt werden. Im zweiten Schritt sollte jedes Unternehmen individuell prüfen, ob weitere Schulungen notwendig sind. Zu empfehlen ist in jedem Fall die Einführung eines KI-Governance-Rahmens mit schriftlichen Leitlinien, welcher als Orientierungshilfe für die Mitarbeitenden dient. Ein KI-Beauftragter ist nach der KI-VO im Gegensatz zur Sicherstellung der KI-Kompetenz nicht verpflichtend. Je nach Unternehmensgröße und Umfang von KI-Nutzung kann dies jedoch sinnvoll sein, um die gesetzlichen Vorgaben effektiv umsetzen zu können.

Gemeinsam mit Ihnen möchten wir die Anforderungen an die KI-Kompetenz aus Art. 4 KI-VO in unserem Webinar am 30. Januar 2025 vom 18-19 Uhr diskutieren, rechtlich und dank unserer Kooperation mit Marcel Pesch von der academy4.ai auch technisch operativ. Marcel Pesch wird uns dann auch das konkrete Schulungsangebot

der Academy4.ai vorstellen, bei dem wir die rechtlichen Schulungsinhalte bereitstellen. Mehr Informationen dazu finden Sie unter https://loschelder.de/de/webinare.html.



KI und Datenschutz: EDSA-Stellungnahme zur Verarbeitung personenbezogener Daten in KI-Modellen

Am 18. Dezember 2024 hat der Europäische Datenschutzausschuss (EDSA) eine neue Stellungnahme zu bestimmten datenschutzrechtlichen Aspekten bei der Verarbeitung von personenbezogenen Daten im Zusammenhang mit KI-Modellen veröffentlicht. Die Stellungnahme enthält richtungsweisende Leitlinien für Entwickler und Nutzer von KI-Modellen und adressiert zentrale Themen wie die Frage, wann ein KI-Modell personenbezogene Daten enthält, die Bewertung des berechtigten Interesses als Rechtsgrundlage für die Entwicklung und Nutzung von KI-Modellen und Rechtsfolgen bei unrechtmäßiger Datenverarbeitung bei Entwicklung eines KI-Modells.

Mit der Stellungnahme 28/2024 vom 17. Dezember 2024 reagiert der EDSA auf eine Anfrage der irischen Aufsichtsbehörde (DPC), die den Datenschutzausschuss gebeten hatte, eine Stellungnahme nach Art. 64 Abs. 2 DSGVO mit allgemeiner Geltung zu erlassen. Im Zusammenhang mit der rasanten Entwicklung von Technologien zur Künstlichen Intelligenz („KI“) kam es in letzter Zeit immer häufiger zu datenschutzrechtlichen Fragestellungen – nicht zuletzt ausgelöst durch den Umstand, dass die im August 2024 in Kraft getretene KI-Verordnung vorsieht, dass die in ihr festgelegten Rechte und

Pflichten neben der DSGVO Anwendung finden (Art. 2 Abs. 7 KI-VO).

Die Stellungnahme des EDSA bezieht sich auf KI-Modelle im Sinne der KI-VO, die das Ergebnis eines Trainings mit personenbezogenen Daten sind. Sie setzt klare Vorgaben, die ein Gleichgewicht zwischen Innovation und Datenschutz schaffen sollen.

Die wichtigsten Punkte der Stellungnahme haben wir hier für Sie zusammengefasst:

Wann ist ein KI-Modell „anonym"?

Für alle Informationen, die sich auf eine identifizierte oder identifizierbare Person beziehen, gelten die Grundsätze der DSGVO. Das gilt auch für pseudonymisierte Daten.

Liegen Daten vor, die sich nicht auf eine identifizierte oder identifizierbare Person beziehen (wie beispielsweise reine Firmendaten, z. B. Geschäftszahlen), müssen die Anforderungen der DSGVO nicht eingehalten werden. Das gleiche gilt bei anonymisierten Daten, die eine Verknüpfung zu einer bestimmten Person nicht (mehr) zulassen.

Sind KI-Modelle trainiert, beinhalten sie selbst i.d.R. keine personenbezogenen Daten in Klarform mehr: Sie beinhalten abstrahierte Datenrepräsentationen, die in Form von Vektoren oder numerischen Merkmalen im Modell vorliegen. Auch während der Anwendung arbeiten KI-Modelle meist auf einer höheren Abstraktionsebene: die Modelle analysieren die eingegeben Daten, indem sie Muster erkennen und Wahrscheinlichkeiten berechnen, um Output zu generieren. Dieser Umstand kann den Anschein erwecken, dass die in dem Model, insbesondere den großen Sprachmodellen (LLM) enthaltenen Datensätze anonym und nicht personenbezogen sind, da sie rein aus der Beziehung und den Zusammenhängen zwischen Token (Wortbausteinen) entstehen.

Allerdings können Daten, die in ein KI-Modell zu Trainingszwecken einfließen, personenbezogen sein und ebenso die Ergebnisse nach einem bestimmten Prompt. Wenn ich ein KI-Modelle nutze, kann ich mithin durch das Prompting personenbezogene Daten „generieren". Sind dann nicht die in einem solchen KI-Modell auch in Ruhe vorliegenden Informationen zwangsläufig auch personenbezogene Daten?

Diesem Punkt widmet sich die erste Frage der DPC, die wissen möchte, ob das anwendungsfähige KI-Modell („final"), das mit personenbezogenen Daten trainiert wurde, den Anforderungen der Definition personenbezogener Daten gem. Art. 4 Abs. 1 DSGVO entspricht („Is the final AI Model, which has been trained using personal data, in all cases, considered not to meet the definition of personal data (as set out in Article 4(1) GDPR)?").

Die Hamburgische Beauftragte für den Datenschutz und die Informationsfreiheit (HmbBfDI) hat im vergangenen Jahr in einem Diskussionspapier hierzu bereits die These zur Diskussion gestellt, dass die bloße Speicherung eines LLMs keine Verarbeitung personenbezogener Daten i. S. d. Art. 4 Nr. 2 DSGVO darstelle. Denn in LLMs selbst seien Tokens, Gewichte und Vektoren, aber keine personenbezogenen Daten gespeichert. Wir berichteten.

Der EDSA sieht das anders und stellt zunächst klar: Bei KI-Modellen, die speziell darauf ausgelegt sind, personenbezogene Daten derjenigen bereitzustellen, deren Daten für das Training genutzt wurden, stelle sich die Frage nach dem Personenbezug nicht; er sei zu bejahen. Beispiele seien generative Sprachmodelle, die mit Sprachaufnahmen einer Person feinabgestimmt würden, um ihre Stimme zu imitieren oder KI-Modelle, die personenbezogen Daten auf Abruf aus dem Training bereitstellten – bei diesen KI-Modellen sei die DSGVO wie gewohnt zu beachten (Rn. 29).

Zur Beantwortung der ersten Frage konzentriert sich der EDSA im Weiteren auf KI-Modelle, die **nicht** darauf ausgelegt sind, personenbezogene Daten aus speziellen Trainingsdaten bereitzustellen. Dabei ist er der Ansicht, dass Informationen aus dem Trainingssatz (einschließlich personenbezogener Daten), in den Parametern des KI-Modells aufgenommen bleiben könnten („may still remain 'absorbed'") – dargestellt durch mathematische Objekte (Rn. 31). Diese Daten unterschieden sich von den eingegebenen Trainingsdaten, enthielten aber stets die ursprünglichen Informationen, die letztendlich extrahiert oder anderweitig, direkt oder indirekt, aus dem KI-Modell gewonnen werden könnten. Der EDSA schlussfolgert daher: Wann immer Informationen, die sich auf identifizierte oder identifizierbare Personen beziehen, deren personenbezogene Daten zum Trainieren des KI-Modells verwendet wurden, mit vernünftigen Mitteln aus dem Modell heraus gewonnen werden können, kann davon ausgegangen werden, dass dieses KI-Modell nicht anonym ist – mit anderen Worten: personenbezogene

Daten enthält. Ob dies tatsächlich so ist, müsse von Fall zu Fall beurteilt werden (Rn. 34).

Damit ein KI-Modell als anonym angesehen werden kann, müssten nach EDSA zwei Voraussetzungen erfüllt sein: (1) Die Wahrscheinlichkeit einer *direkten* Extraktion personenbezogener Daten, die bei der Entwicklung des Modells verwendet wurden, und (2) die Wahrscheinlichkeit, personenbezogene Daten – absichtlich oder unabsichtlich – aus Abfragen zu erhalten, müssen – unter Berücksichtigung aller Mittel, die vernünftigerweise vom Verantwortlichen oder einer anderen Person eingesetzt werden können – gering („insignificant") sein (Rn. 43). Damit bezieht sich der EDSA letztlich auf die Bestimmung relativ anonymer Daten (Erwägungsgrund 26 DSGVO).

Um die Anonymität eines KI-Modells zu überprüfen, sollen die Aufsichtsbehörden die vom Verantwortlichen bereitgestellten Unterlagen prüfen. In der Stellungnahme des EDSA sind Methoden aufgelistet, die von für die Verarbeitung Verantwortlichen für den Nachweis der Anonymität verwendet werden können. Zum Beispiel Methoden, die Verantwortliche in der Entwicklungsphase ergreifen können, um die Sammlung personenbezogener Daten für das Training zu begrenzen oder zu verhindern, die Identifizierbarkeit der genutzten Daten zu verringern oder den Schutz der Daten nach dem Stand der Technik zu gewährleisten (Rn. 44 ff.).

Unter welchen Voraussetzungen können berechtigte Interessen als Rechtsgrundlage für die Entwicklung eines KI-Modells dienen?

Werden personenbezogene Daten verarbeitet, ist stets eine Rechtsgrundlage erforderlich, welche die Verarbeitung gestattet. Diese Rechtsgrundlagen sind unter anderem in Art. 6 DSGVO aufgelistet. Eine dieser Rechtsgrundlagen ist das „berechtigte Interesse"; dieses erlaubt es Unternehmen, personenbezogene Daten zu verarbeiten, wenn die Verarbeitung zur Wahrung der berechtigten Interessen des Verantwortlichen oder eines Dritten erforderlich ist, sofern nicht die Interessen der betroffenen Person überwiegen (Art. 6 Abs. 1 lit. f) DSGVO).

In Bezug auf die zweite und dritte Frage der DPC widmet sich der EDSA in seiner Stellungnahme dem „berechtigten Interesse" als geeignete Rechtsgrundlage für die Verarbeitung personenbezogener Daten bei Entwicklung und Einsatz von KI-Modellen.

Er stellt klar, dass Unternehmen diese Rechtsgrundlage durchaus für die Verarbeitung von personenbezogenen Daten im Zusammenhang mit KI-Modellen heranziehen können, weist jedoch darauf hin, dass der bekannte Drei-Stufen-Test angewendet werden müsse (EDSA, Guidelines 01/24). Welche drei Kriterien nach EDSA kumulativ erfüllt sein müssen, um von einem berechtigten Interesse des Verantwortlichen ausgehen zu können, haben wir hier schon einmal dargestellt: Legitime Ziele, Erforderlichkeit der konkreten Verarbeitung zur Zielerreichung und eine Interessenabwägung im Einzelnen.

Mit Blick auf die Verarbeitung personenbezogener Daten bei Entwicklung und Einsatz von KI-Modellen spielt laut EDSA in der Abwägung insbesondere eine Rolle, welche berechtigten Erwartungen eine betroffene Person im Hinblick auf die Datenverarbeitung haben darf. Für sie kann es schwierig sein, die Vielfalt potentieller Anwendungsmöglichkeiten und Verarbeitungsaktivitäten zu erfassen und die Komplexität der Arbeitsweise eines KI-Modells zu verstehen. Um beurteilen zu können, ob die betroffene Person vernünftigerweise erwarten kann, dass ihre personenbezogenen Daten verarbeitet werden, sind daher die diesen bereitgestellten Informationen als auch der Kontext der Verarbeitung von Bedeutung: Sind die personenbezogenen Daten öffentlich zugänglich oder nicht? In welcher Beziehung steht die betroffene Person zu dem Verantwortlichen der Datenverarbeitung? Aus welcher Quelle stammen die durch das KI-Modell verarbeiteten Daten? Wie wird das KI-Modell möglicherweise in Zukunft weiterverwendet?

Der EDSA schlägt in seiner Stellungnahme außerdem eine Reihe milderer Maßnahmen für die Entwicklungs- und Einsatzphase eines KI-Modells vor, um die Auswirkungen der Datenverarbeitung auf betroffene Personen zu begrenzen (Rn. 96 ff.). Diese sollten auf die Umstände des Falls und die Merkmale des KI-Modells, einschließlich seiner beabsichtigten Nutzung, zugeschnitten sein. Hierunter fallen etwa technische Maßnahmen, z. B. das Ersetzen personenbezogener Daten durch Fake-Daten (insbesondere im Rahmen des LLM-Trainings), und Transparenzmaßnahmen, wie die Veröffentlichung leicht zugänglicher Informationen – über jene nach Art. 13, 14 DSGVO hinaus – zu Kriterien der Datenerhebung und verwendeten Datensätzen.

Welche Folgen hat die unrechtmäßige Datenverarbeitung in der Entwicklungsphase des KI-Modells für spätere Verarbeitungsschritte?

Zu guter Letzt widmet sich der EDSA den Folgen einer unrechtmäßigen Verarbeitung personenbezogener Daten im Rahmen der Entwicklungs- und Trainingsphase für die Rechtmäßigkeit der nachfolgenden Verarbeitung bei Einsatz des KI-Modells. Zur Beantwortung unterscheidet er drei verschiedene Szenarien:

In Szenario 1 werden (nicht anonymisierte) personenbezogene Daten im KI-Model gespeichert und durch denselben Verantwortlichen weiterverarbeitet. Hier ist im Einzelfall zu entscheiden, ob Entwicklungs- und anschließende Bereitstellungsphase unterschiedliche Zwecke verfolgen (und damit unterschiedliche Verarbeitungstätigkeiten darstellen) und inwieweit die Unrechtmäßigkeit der ersten Verarbeitung die Rechtmäßigkeit der nachfolgenden beeinflusst. Beruht die nachfolgende Verarbeitung auf einem berechtigten Interesse gem. Art. 6 Abs. 1 lit. f DSGVO, ist die mögliche Unrechtmäßigkeit einer vorangegangenen Verarbeitung im Rahmen der Abwägungsentscheidung zu berücksichtigen (z.B. mit Blick auf die Risiken für die betroffene Person oder die Tatsache, dass sie mit einer späteren Verarbeitung nicht rechnen musste). In solchen Fällen kann die Unrechtmäßigkeit der ursprünglichen Datenverarbeitung in der Entwicklungsphase Einfluss auf die Rechtmäßigkeit der nachfolgenden Verarbeitung bei Einsatz des KI-Modells haben (Rn. 122, 123).

In Szenario 2 werden personenbezogene Daten in einem KI-Modell durch einen Verantwortlichen gespeichert (auch hier nicht anonymisiert) und durch einen anderen Verantwortlichen – bei Bereitstellung des Modells – weiterverarbeitet. Nach EDSA sollte Letzterer eine angemessene Bewertung durchführen, um sicherzustellen, dass das KI-Modell nicht durch eine unrechtmäßige Verarbeitung personenbezogenen Daten entwickelt wurde (Rn. 129 ff.)

In Szenario 3 werden personenbezogene Daten unrechtmäßig verarbeitet, jedoch werden diese anschließend anonymisiert, bevor sie durch denselben oder einen anderen Verantwortlichen weiterverarbeitet werden. Wenn nachgewiesen werden kann, dass der nachfolgende Einsatz eines KI-Modells keine Verarbeitung

personenbezogener Daten beinhaltet, ist die DSGVO laut EDSA nicht anwendbar und die Unrechtmäßigkeit der anfänglichen Verarbeitung hat keine Auswirkungen auf die nachfolgende Nutzung des KI-Modells (Rn. 134). Wenn ein Verantwortlicher im Anschluss an die Anonymisierung eines KI-Modells erneut personenbezogene Daten bei Einsatz des KI-Modells erhebt, gilt in Bezug auf diese Erhebung und Verarbeitung wiederum die DSGVO. In diesen Fällen beeinflusst die ursprüngliche Unrechtmäßigkeit einer Datenverarbeitung in der Entwicklungsphase die spätere Datenverarbeitung ebenfalls nicht (Rn. 135). Dies wird für die Praxis ein wesentliches Szenario werden, wenn etwa Dritte die von anderen entwickelten Modelle weiterverwenden und so eine „Vergiftung" der nachfolgenden Verarbeitungsprozesse vermeiden können (Vermeidung „fruit of the poisonous tree").

Die Stellungnahme des EDSA ist ein erster Wegweiser, der die spezifischen Herausforderungen des Datenschutzes, die KI-Modelle mit sich bringen, auf europäischer Ebene adressiert und mehr Klarheit in Bereichen schafft, die bislang von Unsicherheit geprägt waren. Wir werden in diesem Jahr noch intensiv über KI und Datenschutz diskutieren – vielleicht ja gemeinsam in Berlin am 21. Februar 2025 auf dem DAV KI-Forum, in dem unsere Partnerin Dr. Kristina Schreiber unter Moderation von Frederik Richter, LL.M., Stiftung Datenschutz, auf dem Panel mitdiskutieren wird (https://anwaltverein.de/de/der-dav/dav-veranstaltungen/dav-forum/dav-ki-forum).

KI und Datenschutz: Large Language Models und Data Sharing

Die Internationale Arbeitsgruppe für Datenschutz in der Technologie (IWGDPT) – auch genannt „Berlin Group" – hat am 27. Dezember 2024 unter dem Vorsitz der Bundesbeauftragten für Datenschutz und Informationsfreiheit zwei wegweisende Arbeitspapiere zum Thema "Large Language Models" (LLMs) und "Data Sharing" veröffentlicht.

Um Innovationen voranzutreiben, neue Erkenntnisse zu gewinnen und neue Produkte und Dienstleistungen zu entwickeln, kommt es immer häufiger zu „Data Sharing" zwischen Organisationen. Darunter wird die Bereitstellung von Daten zur Nutzung durch andere verstanden. Es findet also eine Datenübertragung zwischen Organisationen statt. Unter Umständen kann das Teilen von Daten bzw. der Austausch von Daten, beispielsweise aufgrund mangelnder Transparenz oder fehlender Sicherheitsvorkehrungen, zu größeren Datenschutzverstößen und Datenmissbrauch führen. Die Berlin-Group hat sich dieser Problematik in ihrem Papier zum „Data-Sharing" angenommen und Empfehlungen bzw. Leitlinien für die verschiedenen Akteure des Datenaustausch entwickelt, um eine vertrauenswürdige Umgebung für Datenaustausch zu schaffen. Dabei sollen Datenschutzgrundsätze gewahrt bleiben und gleichzeitig das Potenzial eines sicheren und geschützten Datenaustauschs maximiert werden.

Zur Generierung von Texten werden oftmals sog. Large Language Models (kurz LLMs) eingesetzt. Damit das gelingt werden diese vorab mit großen Datenmengen trainiert – sind darunter personenbezogene Daten, kann es unter Umständen auch hier zu Datenschutzverstößen kommen. In ihrem zweiten Papier zu „Large Language Models" gibt die Berlin Group einen umfassenden Überblick darüber, wie ein datenschutzkonformer Einsatz derer gelingen kann.

Data-Sharing

Weltweit gilt es bereits zahlreiche gesetzgeberische Initiativen, die das Data Sharing im öffentlichen als auch im privaten Sektor regulieren. Beispiele hierfür sind der Data Governance Act und der Data Act auf EU-Ebene, die Regeln für den Austausch von Daten festlegen und den Zugang zu Daten ermöglichen. Dennoch birgt der

Austausch von Daten immer das Risiko von Datenmissbrauch und Datenschutzverstößen.

Um das zu vermeiden, wird in dem Papier zum einen empfohlen, Privacy-Enhancing Technologies (PETs) zu implementieren. Um Datenschutzvorgaben einzuhalten, sollte zudem eine exakte Analyse darüber stattfinden, warum und an wen Daten geschickt werden. Zudem enthält das Papier Empfehlungen für verschiedene Akteure wie Datenverantwortliche, Gesetzgeber, Technologieanbieter, Forschungsgemeinschaft und Datenschutzbehörden. Betroffene sollten ausreichend über den Datenaustausch informiert werden und müssen die Möglichkeit haben einer solchen Datenübermittlung widersprechen zu können.

Large-Language-Models

Künstliche Intelligenz, die Texte in menschlicher Sprache generieren kann, basieren meist auf einem LLM. Die Grundlage für diese KI-Modelle bilden meist große Datenmengen, mit denen die LLMs vorab trainiert werden. In dem Papier der Berlin Group zu diesen KI-Modellen werden die Risiken solcher KI-Modelle für den Datenschutz aufgezeigt. So kann es durch die Nutzung von LLMs beispielsweise dazu kommen, dass Desinformationen verbreitet werden, da die LLMs fehlerhafte Inhalte generieren oder Betroffene die Kontrolle über ihre Daten verlieren, da unklar ist, inwiefern die Daten genutzt werden.

Um das zu vermeiden, betont die Arbeitsgruppe immer wieder die Bedeutung hochwertiger und datenschutzkonformer Trainingsdaten. Zudem soll neben den allgemeinen Prinzipien des Datenschutzes wie Transparenz, Zweckbegrenzung und Erlaubnisgrundlage für die Datenverarbeitung, auch daran gearbeitet werden, die Verarbeitung von personenbezogenen Daten durch LLMs möglichst zu vermeiden.

Gleichzeitig werden technische Designs der KI-Modelle vorgeschlagen, um Datenschutzgrundsätze technisch zu ermöglichen. Durch bspw. Differential Privacy, kann die Wahrscheinlichkeit reduziert werden, dass personenbezogene Daten rekonstruiert werden oder durch Machine Unlearning, welches es erlaubt, Daten nachträglich aus einem Modell zu entfernen. Dadurch wird deutlich, dass der erfolgreiche und datenschutzkonforme

Einsatz von LLMs auch innovative technische Maßnahmen gewährleistet werden.

VW Datenpanne: Zahlreiche Standortdaten im Internet zugänglich

Dass Autos heutzutage mit einer eigenen Software ausgestattet und mit Apps verknüpft sind, ist mittlerweile der Normalfall. Diese Anwendungen ermöglichen oft erst die Ausschöpfung der vollen Funktionen des Fahrzeugs. Zudem können damit Daten über das Fahrzeug und dessen Nutzung gesammelt werden, was von den Automobilunternehmen umfassend genutzt wird, nicht zuletzt für die Weiterentwicklung der Fahrzeuge und ihrer Funktionen. Dass Softwarehersteller hierbei besonders auf die Sicherheit dieser Daten achten müssen, verdeutlichte jüngst eine Datenpanne bei Volkswagen.

800.000 Fahrzeugdaten online zugänglich

„Kurios bis bedenklich" ordnete der Chaos Computer Club (CCC) seine Entdeckung ein, die er zu einem wortwörtlich bewegenden Abschluss des Jahres 2024 mit der (Datenschutz)Welt teilte: Bewegungsdaten von ca. 800.000 Fahrzeugen waren über mehrere Monate im Internet zugänglich. Die Daten lagen in einem Cloud-Speicher, der mit den nötigen Programmen und IT-Wissen relativ leicht aufzufinden war. Betroffen waren gewisse Modelle von Elektrofahrzeugen der Marken VW, Skoda, Audi und Seat. Einsehbar

waren Daten über den Batterieladestand, den Inspektionsstatus oder darüber, ob der Motor gerade eingeschaltet war.

Doch das ist bei Weitem nicht alles: Bei 460.000 Fahrzeugen – und damit mehr als der Hälfte – waren die exakten Positionen der Abstellorte der Fahrzeuge mit Uhrzeiten einsehbar. Die Standorte wurden durch Längen- und Breitengrade angegeben und waren bei VW- und Seat-Modellen bis auf zehn Zentimeter (!) genau. Die Standorte der anderen Pkws waren bis auf zehn Kilometer genau, was deutlich weniger Aussagekraft hat, an sich jedoch nicht weniger problematisch ist. Denn bei den meisten Fahrzeugdaten konnte über die ebenfalls einsehbaren Zugangsdaten der Nutzer der VW-App eine Verknüpfung zu Namen und Kontaktdaten der Fahrer oder Eigentümer hergestellt werden. So ließen die in dem Cloud-Speicher belegenden Datensätze Einblicke in den Alltag und die Bewegungsmuster etlicher Menschen zu.

Die meisten Daten betreffen Fahrzeuge in Deutschland. Aber auch Daten von Fahrzeugen in Norwegen, Schweden, Großbritannien und den Niederlanden tauchten in dem Cloud-Speicher auf.

Zügige Reaktion vermeidet unter Umständen schwere Folgen

Verantwortlich für diese Veröffentlichung war die VW-Tochtergesellschaft Cariad, zuständig für die Software-Entwicklung der VW-Group, die jedoch nicht selbst auf ihren Fehler aufmerksam wurde. Ein anonymer Hinweisgeber wandte sich an den CCC und den Spiegel.

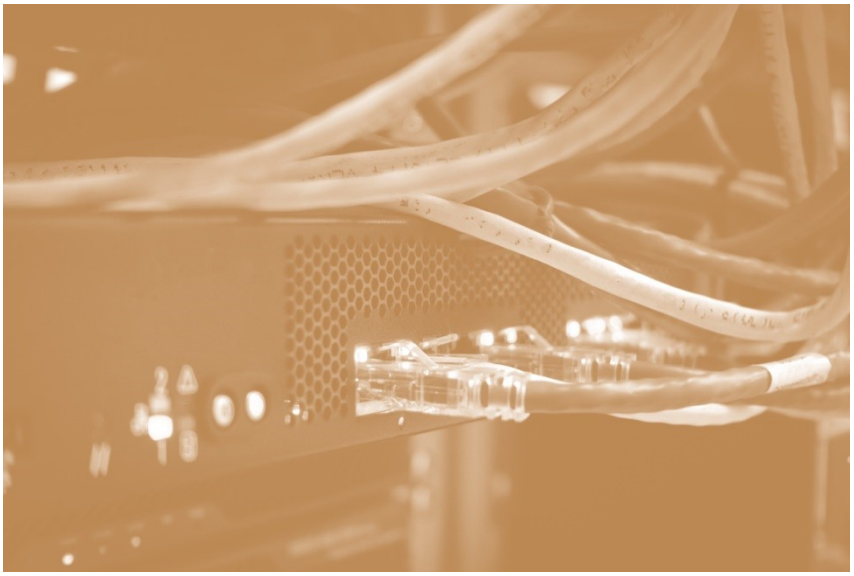
Den erheblichen Gefahren, die aus einer solchen „Fehlkonfiguration“, wie Cariad sie nannte, folgen können, sollten sich Softwareentwickler besonders bewusst sein. Das zeigt diese Sicherheitslücke ganz besonders: Sie betraf Fahrzeug- und Bewegungsdaten zahlreicher Politiker, Polizeibeamter und Nachrichtendienstmitarbeiter. In solchen Fällen sind genaue Aufenthaltsorte oft sehr sensible Informationen. Aber auch für Privatpersonen kann es durchaus gefährlich werden, wenn einsehbar ist, wann sie zu Hause sind oder wann nicht. Mit Blick auf Einbrüche und Erpressungen bieten diese sensiblen Angaben eine wahre Datenschatzgrube für Kriminelle! Auch glaubwürdige Phishingmails par excellence hätten mithilfe dieser Informationen erstellt werden können, um an Kreditkarten und Zahlungsinformationen von Kunden zu gelangen.

Cariad hat zügig reagiert und die Verantwortung übernommen. Die Sicherheitslücke wurde geschlossen. Auch für betroffene Personen gab es Entwarnung: Das Unternehmen informierte darüber, dass es unwahrscheinlich sei, dass irgendjemand außer den Sicherheitsexperten, die das Datenleck untersucht hatten, tatsächlich auf die Daten zugegriffen habe. Auch der CCC habe lediglich auf Daten zugreifen können, die keine Rückschlüsse auf einzelne Personen zuließen.

Dennoch gilt: Solche Vorkommnisse sollten im Vorfeld vermieden werden, denn Vorsicht ist besser als Nachsicht – ganz besonders, wenn es um die IT- und Datensicherheit geht.

Verschärfung der IT-Sicherheit

Die IT-Sicherheit wird derzeit auch regulatorisch für viele zunehmend verschärft: Die NIS-2-Richtlinie ist zeitnah in nationales Recht umzusetzen und verpflichtet rund 30.000 Unternehmen in Deutschland zu einer verschärften IT-Sicherheit. Ende 2024 wurde zudem der Cyber Resilience Act, eine EU-Verordnung, verkündet, die ab 2027 zu umfassenden IT-Sicherheitsvorgaben für digitale Produkte von der Software bis zum IoT-Gerät verpflichtet. Unter diesen neuen Rechtsakten ist eine Panne wie bei VW dann neben dem Datenschutzrecht auch nach diesen Rechtsakten strafbewehrt und abzustellen.



Zu guter Letzt

Wer ab und zu genauer hinschaut, stößt gegebenenfalls auf Lücken in der Datenschutzpolitik einiger Unternehmen. Aus verschiedenen Anlässen erfolgten in den letzten Jahren einige datenschutzrechtliche Prüfungen unter anderem bei Apple, Netflix und Coolblue, die zuletzt jeweils hohe Bußgelder nach sich zogen. Dabei ging es um das Mithören von Privatgesprächen, die Informationspflicht nach der DSGVO und die Einwilligung in die Nutzung von Cookies auf einer Website.

- **Apple stimmt Zahlung von 95 Millionen Dollar wegen der Aufnahme privater Gespräche durch Siri zu**

Nicht selten wundert man sich über Werbung auf Social Media, die genau zu den Themen passt, über die man sich gerade erst – vermeintlich privat – unterhalten hat. Dieses Phänomen hat 2019 zahlreiche Nutzer von Apple-Geräten dazu veranlasst, Sammelklage gegen das Unternehmen zu erheben. Der Vorwurf: Der Sprachassistent Siri habe private Gespräche abgehört, ohne dass die Nutzer Siri ausdrücklich mit den Worten „Hey Siri“ aktiviert hätten. Diese Gespräche seien darüber hinaus mit Dritten geteilt und dazu verwendet worden, individuell angepasste Werbung zu zeigen.

Apple hat nun in einem Vergleich zugestimmt, wegen des mutmaßlichen Mithörens von Privatgesprächen mit Apple-Geräten 95 Millionen Dollar an betroffene Nutzer zu zahlen. In der Vereinbarung hat Apple außerdem zugesichert, die rechtswidrig registrierten Gespräche zu löschen. Zudem sollen die Nutzer von Apple-Geräten mehr Einstellungsoptionen zur Aufzeichnung ihrer Stimmen durch Siri bekommen. Der Vergleich soll endgültig am 14. Februar 2025 gerichtlich genehmigt werden. Ein eigenes Fehlverhalten weist Apple jedoch weiterhin konsequent zurück. Bei Erfolg der Sammelklage ohne den jetzt geschlossenen Vergleich, hätte die Strafe nach US-amerikanischem Recht nämlich voraussichtlich über 1,5 Milliarden Dollar betragen. Der Konzern hat sich für den Vergleich entschieden, um weitere Kosten und Unsicherheiten – auch für die betroffenen Nutzer – zu vermeiden.

- **Bußgeld gegen Netflix wegen ungenügender Informationen über Datenverarbeitungen**

Netflix wird in zahlreichen Haushalten tagtäglich genutzt. Dabei sammelt der Streaming-Dienst viele personenbezogene Daten seiner Nutzer, wie E-Mail-Adressen, Telefonnummern oder Zahlungsdaten. Wer personenbezogene Daten erhebt, ist laut DSGVO verpflichtet, die betroffenen Personen darüber zu informieren, warum welche Daten gesammelt werden und wie diese verarbeitet werden.

Die niederländische Datenschutzbehörde beurteilte die durch Netflix bereitgestellten Informationen jetzt als zu unklar und unzureichend. Auch seien datenschutzrechtliche Auskunftsverlangen (Art. 15 DSGVO) nicht ausreichend beantwortet worden. Unklarheiten bestanden nach Ansicht der Datenschutzbehörde insbesondere hinsichtlich der Verarbeitungszwecke, der Rechtsgrundlage zur Datenverarbeitung, der Art der verarbeiteten Daten, der Dauer ihrer Speicherung und der Sicherheitsmaßnahmen bei Datenweitergabe in Staaten außerhalb der EU. Wegen dieser Verstöße verhängte sie gegen Netflix ein Bußgeld in Höhe von 4,75 Millionen Euro. Anstoß für die Überprüfung hatte bereits vor fünf Jahren die österreichische NGO None of your business (noyb) gegeben. Und es bleibt weiterhin spannend: Netflix hat gegen das Bußgeld Einspruch erhoben. Die Datenschutzrichtlinie des Unternehmens und die erteilten Informationen wurden mittlerweile aber aktualisiert.

- **Fehlende Cookie-Einwilligung führt zu 40.000 Euro Bußgeld für Coolblue**

Neuigkeiten gibt es von der niederländischen Datenschutzbehörde gleich in zwei Fällen: Auch dem Technik-Onlineshop Coolblue hat sie ein – deutlich geringeres und dennoch eindrückliches – Bußgeld von 40.000 Euro auferlegt. Hintergrund ist der rechtswidrige Einsatz von Cookies im Jahr 2020. Die DSGVO verlangt für die Verarbeitung personenbezogener Daten auf Webseiten durch den Einsatz von Cookies in vielen Fällen die aktive Einwilligung der Website-Besucher. Coolblue hatte sich jedoch auf die Annahme verlassen, die betroffenen Personen erteilten ihre Einwilligung automatisch bereits durch den bloßen Besuch der Website. Statt die Cookie-Einwilligung der Besucher aktiv durch ein sog. Opt-in einzuholen, setzte Coolblue entsprechende Haken zur Cookie-Zustimmung selbst.

Nach einer behördlichen Überprüfung der niederländischen Coolblue-Website im Jahr 2019 wurde das E-Commerce-Unternehmen auf den DSGVO-Verstoß aufmerksam gemacht und aufgefordert, seine Website an die datenschutzrechtlichen Vorgaben anzupassen. Die Anpassung der Cookie-Einstellungen nahm Coolblue – nach erneuter Aufforderung durch die Datenschutzbehörde – schließlich im Juni 2020 vor.

Für alle weiteren Fragen rund um das Datenschutzrecht stehen Ihnen gerne zur Verfügung

Dr. Kristina Schreiber
+49 221 65065-337
kristina.schreiber@loschelder.de

Dr. Simon Kohm
+49 221 65065-200
simon.kohm@loschelder.de

Dennis Pethke, LL.M.
+49 221 65065-337
dennis.pethke@loschelder.de

Rebecca Moßner
+49 221 65065-337
rebecca.mossner@loschelder.de

Impressum

LOSCHELDER RECHTSANWÄLTE
Partnerschaftsgesellschaft mbB
Konrad-Adenauer-Ufer 11
50668 Köln

Tel. +49 (0)221 65065-0, Fax +49 (0)221 65065-110
info@loschelder.de
www.loschelder.de