Stellungnahme des EDSA nach Artikel 64 DSGVO

Stellungnahme 28/2024 zu gewissen Datenschutzaspekten der Verarbeitung personenbezogener Daten im Zusammenhang mit KI-Modellen

Angenommen am 2. Dezember 2024

Im englischen Original wird der 17.12.2024 genannt.]

Translations proofread by EDPB Members.

This language version has not yet been proofread.

Zusammenfassung

In einem breiten Spektrum von Wirtschaftssektoren und gesellschaftlichen Aktivitäten können KI-Technologien vielfältige Chancen eröffnen und Nutzen bieten.

Indem die DSGVO das Grundrecht auf Datenschutz schützt, werden nicht nur diese Chancen durch die DSGVO geschützt, sondern auch andere Unionsgrundrechte gewahrt, unter anderem die Rechte auf Gedanken-, Meinungsäußerungs- und Informationsfreiheit, auf Bildungs- und auf unternehmerische Freiheit. Auf diese Weise bietet die DSGVO einen Rechtsrahmen, der verantwortungsvolle Innovation fördert.

Im Hinblick auf die Datenschutzfragen, die durch diese Technologien aufgeworfen werden, hat die irische Aufsichtsbehörde den EDSA in diesem Zusammenhang gemäß Artikel 64 Absatz 2 DSGVO um eine Stellungnahme zu einer Angelegenheit mit allgemeiner Geltung ersucht. Das Ersuchen betrifft die Verarbeitung personenbezogener Daten im Zuge der Entwicklungsphase und der Einsatzphase von Modellen für künstliche Intelligenz („**KI**"). Im Einzelnen wurde mit dem Ersuchen gefragt: 1) wann und unter welchen Voraussetzungen ein KI-Modell für „anonym" befunden werden kann, 2) auf welche Weise Verantwortliche die Angemessenheit des berechtigten Interesses als eine Rechtsgrundlage nachweisen können, und zwar in der Entwicklungsphase und 3) in der Einsatzphase, 4) und welche Folgen eine rechtswidrige Verarbeitung personenbezogener Daten in der Entwicklungsphase eines KI-Modells auf die spätere Verarbeitung oder den späteren Betrieb des KI-Modells hat.

In Bezug auf die erste Frage heißt es in der Stellungnahme, dass die geltend gemachte Anonymität eines KI-Modells von den zuständigen Aufsichtsbehörden auf Einzelfallbasis geprüft werden sollten, da nach Ansicht des EDSA KI-Modelle, die mit personenbezogenen Daten trainiert wurden, nicht in allen Fällen für anonym befunden werden können. Ein KI-Modell kann nur dann für anonym befunden werden, wenn zwei kumulative Voraussetzungen erfüllt sind: 1) Sowohl die Wahrscheinlichkeit direkter (einschließlich probabilistischer) Entnahmen personenbezogener Daten über die natürlichen Personen, deren personenbezogene Daten für die Modellentwicklung verwendet wurden, 2) als auch die Wahrscheinlichkeit, dass solche personenbezogenen Daten, ob vorsätzlich oder nicht vorsätzlich, durch Anfragen erlangt werden, sollte bei Berücksichtigung *„alle[r] Mittel …, die von dem Verantwortlichen oder einer anderen Person nach allgemeinem Ermessen wahrscheinlich genutzt werden*" vernachlässigbar gering sein.

Zur Vornahme ihrer Bewertung sollten die Aufsichtsbehörden die vom Verantwortlichen zum Nachweis der Anonymität des Modells vorgelegte Dokumentation prüfen. Diesbezüglich bietet die Stellungnahme eine unverbindliche und nicht erschöpfende Liste der Methoden, derer sich Verantwortliche zum Nachweis der Anonymität bedienen können; diese Liste kann folglich auch von den Aufsichtsbehörden für die Bewertung der von einem Verantwortlichen geltend gemachten Anonymität in Betracht gezogen werden. Darunter fallen beispielsweise die Ansätze, nach denen die Verantwortliche in der Entwicklungsphase die Erhebung für das Training verwendeter personenbezogener Daten verhindern oder einschränken, um deren Identifizierbarkeit zu reduzieren, um deren Entnahme zu verhindern oder um dem Stand der Technik entsprechende Angriffsresistenz zu sicherzustellen.

In Bezug auf die zweite und dritte Frage enthält die Stellungnahme allgemeine Erwägungen, die bei der Bewertung, ob sich Verantwortliche für die im Zuge der Entwicklung und des Einsatzes von KI-Modellen erfolgte Verarbeitung auf berechtigte Interessen als geeignete Rechtsgrundlage stützen können, von den Aufsichtsbehörden zu berücksichtigen sind.

In der Stellungnahme wird daran erinnert, dass es keine Hierarchie der in der DSGVO genannten Rechtsgrundlagen gibt und dass es Sache der Verantwortlichen ist, die für ihre Verarbeitungstätigkeiten geeignete Rechtsgrundlage anzugeben. Sodann wird in der Stellungnahme an die dreistufige Prüfung erinnert, nach der die Bewertung, ob berechtigte Interessen als Rechtsgrundlage in Betracht kommen, vorgenommen werden sollte; danach ist zu prüfen: 1) welche berechtigten Interessen der Verantwortliche oder ein Dritter verfolgt; 2) ob die Verarbeitung personenbezogener Daten zur Verwirklichung eines oder mehrerer der verfolgten berechtigten Interessen erforderlich ist (auch als „Erforderlichkeitsprüfung“ bezeichnet) und 3) ob nicht die Interessen oder Grundrechte und Grundfreiheiten der Personen, deren Daten geschützt werden sollen, gegenüber dem berechtigten Interesse des Verantwortlichen oder eines Dritten überwiegen (auch als „Abwägungsprüfung“ bezeichnet).

In Bezug auf den ersten Prüfungsschritt wird in der Stellungnahme daran erinnert, dass ein Interesse als berechtigt angesehen werden kann, wenn die folgenden drei kumulativen Voraussetzungen erfüllt sind: Das Interesse ist 1) rechtmäßig; es ist 2) klar und deutlich dargelegt; und es ist 3) real und gegenwärtig (d. h. nicht spekulativ). Ein solches Interesse mag für die Entwicklung eines KI-Modells (z. B die Entwicklung eines Konversationsagenten, der Nutzern Hilfe leistet) oder für dessen Einsatz (z. B. zur Verbesserung der Bedrohungserkennung in einem Informationssystem) gegeben sein.

In Bezug auf den zweiten Prüfungsschritt wird in der Stellungnahme daran erinnert, dass für die Erforderlichkeitsprüfung darauf abzustellen ist: 1) ob die Verarbeitungstätigkeit die Verfolgung des berechtigten Interesses ermöglicht; und 2) ob es keinen weniger eingreifenden Weg gibt, dieses Interesse zu verfolgen. Für die Bewertung, ob die Erforderlichkeitsvoraussetzung erfüllt ist, sollten die Aufsichtsbehörden insbesondere auf die Menge der verarbeiteten personenbezogenen Daten achten, sowie darauf, ob diese in angemessenem Verhältnis zu den im Einzelfall verfolgten berechtigten Interessen steht, was auch im Licht des Grundsatzes der Datenminimierung zu beurteilen ist.

In Bezug auf den dritten Prüfungsschritt wird in der Stellungnahme daran erinnert, dass die Abwägungsprüfung unter Berücksichtigung der besonderen Umstände des Einzelfalls vorzunehmen ist. Sodann wird ein Überblick über die Elemente gegeben, die bei der Beurteilung, ob die Interessen oder Grundrechte und Grundfreiheiten der von der Datenverarbeitung betroffenen Personen gegenüber dem Interesse des Verantwortlichen oder Dritten überwiegen, von den Aufsichtsbehörden zu berücksichtigen sind.

Im Zusammenhang mit dem dritten Prüfungsschritt werden in der Stellungnahme die besonderen Risiken für die Grundrechte hervorgehoben, die sich in der Entwicklungsphase bzw. in der Einsatzphase von KI-Modellen ergeben können. Außerdem wird klargestellt, dass die Verarbeitung personenbezogener Daten, die im Zuge der Entwicklungsphase bzw. der Einsatzphase von KI-Modellen erfolgt, unterschiedliche – positive oder negative – Auswirkungen auf die Personen haben kann, die von der Verarbeitung personenbezogener Daten betroffen sind. Bei der Bewertung dieser Auswirkungen können die Aufsichtsbehörden die Art der von den Modellen verarbeiteten Daten, den Kontext der Verarbeitung sowie die möglichen weiteren Folgen der Verarbeitung berücksichtigen.

In der Stellungnahme wird aufgezeigt, welche Bedeutung den vernünftigen Erwartungen der betroffenen Personen bei der Abwägungsprüfung zukommt. Diese kann wichtig sein, zum einen wegen der Komplexität der in KI-Modellen verwendeten Technologien, zum anderen im Hinblick darauf, dass es für betroffene Personen möglicherweise schwierig ist, deren vielfältige Nutzungsmöglichkeiten wie auch die verschiedenen damit verbundenen Verarbeitungsvorgänge zu verstehen. Insoweit können für die Bewertung, ob die betroffenen Personen vernünftigerweise erwarten können, dass ihre personenbezogenen Daten verarbeitet werden, neben anderen Aspekten, sowohl die den betroffenen

Personen gegebenen Informationen wie auch der Kontext der Verarbeitung zu prüfen sein. Was die Umstände angeht, kann Folgendes in Betracht kommen: ob die personenbezogenen Daten öffentlich verfügbar waren, die Art des Verhältnisses zwischen der betroffenen Person und dem Verantwortlichen (und ob zwischen ihnen eine Verbindung besteht), die Art der Dienstleistung, die Umstände, unter denen die personenbezogenen Daten erhoben wurden, die für die Datenerhebung genutzte Quelle (d. h. die Website oder der Dienst, die/der für die Erhebung der personenbezogenen Daten genutzt wurde, und die dort angebotenen Datenschutzeinstellungen), die potenziellen weiteren Verwendungen des Modells sowie ob sich die betroffenen Personen überhaupt dessen bewusst sind, dass ihre personenbezogenen Daten online sind.

In der Stellungnahme wird auch daran erinnert, dass in Fällen, in denen die Interessen, Grundrechte und Grundfreiheiten der betroffenen Personen die vom Verantwortlichen oder einem Dritten verfolgten berechtigten Interessen überwiegen, der Verantwortliche in Betracht ziehen kann, risikomindernde Maßnahmen zu ergreifen, um die Auswirkungen der Verarbeitung auf die betroffenen Personen in Grenzen zu halten. Risikominderungsmaßnahmen dürfen jedoch nicht mit den Maßnahmen verwechselt werden, zu denen der Verantwortliche gesetzlich verpflichtet ist, um die DSGVO-Konformität sicherzustellen. Außerdem sollten die Maßnahmen auf die Umstände des Einzelfalls und die besonderen Eigenschaften des KI-Modells sowie dessen beabsichtigte Verwendung zugeschnitten sein. Diesbezüglich bietet die Stellungnahme eine nicht erschöpfende Liste mit Beispielen für Risikominderungsmaßnahmen in der Entwicklungsphase (auch in Bezug auf das Auslesen von Daten aus Online-Quellen („Web Scraping“)) und in der Einsatzphase. Risikominderungsmaßnahmen unterliegen möglicherweise einer rapiden Weiterentwicklung und sollten auf die Umstände des Einzelfalls zugeschnitten sein. Deshalb ist die Geeignetheit der umgesetzten Risikominderungsmaßnahmen von den Aufsichtsbehörden jeweils auf Einzelfallbasis zu bewerten.

In Bezug auf die vierte Frage wird in der Stellungnahme allgemein daran erinnert, dass den Aufsichtsbehörden sowohl für die Bewertung etwaiger Datenschutzverletzung(en) als auch für die Auswahl geeigneter, notwendiger und verhältnismäßiger Maßnahmen Ermessen eingeräumt ist, wobei allerdings die Umstände des Einzelfalls zu berücksichtigen sind. Sodann werden in der Stellungnahme drei Szenarien betrachtet:

In Szenario 1 verbleiben personenbezogene Daten im KI-Modell (weshalb das Modell, wie oben zur ersten Frage ausgeführt wurde, nicht für anonym befunden werden kann) und diese werden später vom selben Verantwortlichen (beispielsweise im Zusammenhang mit dem Einsatz des Modells) verarbeitet. In der Stellungnahme wird ausgeführt, dass die Fragen, ob in der Entwicklungsphase und in der Einsatzphase jeweils gesonderte Zwecke verfolgt werden (sodass diese gesonderte Verarbeitungsvorgänge darstellen) und inwieweit das Fehlen einer Rechtsgrundlage für die anfängliche Verarbeitungstätigkeit Auswirkungen auf die Rechtmäßigkeit der späteren Verarbeitung hat, jeweils nach den Umständen des Einzelfalls bewertet werden sollten.

In Szenario 2 verbleiben personenbezogene Daten im Modell, die später von einem anderen Verantwortlichen, der das Modell einsetzt, verarbeitet werden. Dazu heißt es in der Stellungnahme, dass die Aufsichtsbehörden berücksichtigen sollten, ob sich der Verantwortliche, der das Modell einsetzt, im Rahmen seiner Rechenschaftspflichten zum Nachweis der Einhaltung von Artikel 5 Absatz 1 Buchstabe a DSGVO und Artikel 6 DSGVO in geeigneter Weise vergewissert hat, dass für die Entwicklung des KI-Modells keine rechtswidrig verarbeiteten personenbezogenen Daten verarbeitet wurden. Für diese Bewertung sollte beispielsweise berücksichtigt werden, aus welcher Quelle die personenbezogenen Daten stammen und ob hinsichtlich der Verarbeitung in der Entwicklungsphase eine Datenschutzverletzung festgestellt wurde (insbesondere wenn es sich um eine

aufsichtsbehördliche oder gerichtliche Feststellung handelte); die Detailliertheit der Bewertung sollte von den durch die Verarbeitung in der Einsatzphase entstehenden Risiken abhängen.

In Szenario 3 verarbeitet ein Verantwortlicher im Zuge der Entwicklung eines KI-Modells rechtswidrig personenbezogene Daten; anschließend stellt dieser Verantwortliche jedoch sicher, dass das KI-Modell anonymisiert wird, bevor dann im Zuge des Modelleinsatzes er selbst oder ein anderer Verantwortlicher eine andere Verarbeitung personenbezogener Daten einleitet. Sofern Nachweis dafür erbracht werden kann, dass der anschließende Betrieb des KI-Modells nicht mit der Verarbeitung personenbezogener Daten einhergeht, wäre die DSGVO laut der Stellungnahme des EDSA nicht anwendbar. Die Rechtswidrigkeit der anfänglichen Verarbeitung sollte also keine Auswirkungen auf den späteren Betrieb des Modells haben. Der EDSA weist aber darauf hin, dass die DSGVO sehr wohl insoweit Anwendung findet, als Verantwortliche personenbezogene Daten verarbeiten, die nach der Anonymisierung des Modells in der Einsatzphase erhoben wurden. Was die DSGVO angeht, sollte – so die Stellungnahme – die Rechtswidrigkeit der anfänglichen Datenverarbeitung die Rechtmäßigkeit der in der Einsatzphase ausgeführten Verarbeitung unberührt lassen.

Inhaltsverzeichnis

Der Europäische Datenschutzausschuss –

gestützt auf Artikel 63 und Artikel 64 Absatz 2 der Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (im Folgenden „DSGVO“),

gestützt auf das EWR-Abkommen, insbesondere auf Anhang XI und das Protokoll 37, in der durch den Beschluss des Gemeinsamen EWR-Ausschusses Nr. 154/2018 vom 6. Juli 2018 geänderten Fassung[1],

gestützt auf Artikel 10 und Artikel 22 seiner Geschäftsordnung –

in Erwägung nachstehender Gründe:

(1) Die Hauptaufgabe des Europäischen Datenschutzausschusses (im Folgenden der „**Ausschuss**“ oder der „**EDSA**“) besteht darin, die kohärente Anwendung der DSGVO im gesamten Europäischen Wirtschaftsraum („**EWR**“) sicherzustellen. Nach Artikel 64 Absatz 2 DSGVO können jede Aufsichtsbehörde („**AB**“), der Vorsitz des Ausschusses oder die Kommission beantragen, dass eine Angelegenheit mit allgemeiner Geltung oder mit Auswirkungen in mehr als einem EWR-Mitgliedstaat vom Ausschuss geprüft wird, um eine Stellungnahme zu erhalten. Diese Stellungnahme dient der Prüfung einer Angelegenheit mit allgemeiner Geltung oder mit Auswirkungen in mehr als einem EWR-Mitgliedstaat.

(2) Die Stellungnahme des Ausschusses wird gemäß Artikel 64 Absatz 3 DSGVO in Verbindung mit Artikel 10 Absatz 2 der Geschäftsordnung des EDSA binnen acht Wochen nach dem Beschluss des Vorsitzes und der zuständigen Aufsichtsbehörde über die Vollständigkeit des Dossiers angenommen. Durch Beschluss des Vorsitzes kann diese Frist unter Berücksichtigung der Komplexität der Angelegenheit um weitere sechs Wochen verlängert werden.

HAT FOLGENDE STELLUNGNAHME ERLASSEN:

1 Einleitung

1.1 Zusammenfassung des Sachverhalts

1. Am 4. September 2024 ersuchte die irische Aufsichtsbehörde (die „**IE-AB**“ oder „**ersuchende AB**“) den EDSA um eine Stellungnahme gemäß Artikel 64 Absatz 2 DSGVO zu KI-Modellen und der Verarbeitung personenbezogener Daten (das „**Ersuchen**“).

2. Am 13. September 2024 erklärten der Vorsitz des Ausschusses und die irische Aufsichtsbehörde das Dossier für vollständig. Am folgenden Arbeitstag, dem 16. September 2024, wurde die Akte dem Sekretariat des EDSA übermittelt. Angesichts der Komplexität der Angelegenheit beschloss der Vorsitz des Ausschusses, die gesetzliche Frist im Einklang mit Artikel 64 Absatz 3 DSGVO und Artikel 10 Absatz 4 der Geschäftsordnung des EDSA zu verlängern.

[1] Soweit in dieser Stellungnahme auf „Mitgliedstaaten“ Bezug genommen wird, ist dies als Bezugnahme auf „EWR-Mitgliedstaaten“ zu verstehen. Soweit in dieser Stellungnahme auf die „Union“ Bezug genommen wird, ist dies als Bezugnahme auf den „EWR“ zu verstehen.

3. Das Ersuchen betrifft gewisse Elemente des Trainings, der Aktualisierung, der Entwicklung und des Betriebs von KI-Modellen, bei denen der relevante Datensatz personenbezogene Daten enthält. Die irische Aufsichtsbehörde weist darauf hin, dass das Ersuchen wichtige Fragen betrifft, die sich in hohem Maße auf die betroffenen Personen und die Verantwortlichen im EWR auswirken, und dass es bislang noch keinen harmonisierten Standpunkt unter den nationalen Aufsichtsbehörden gibt[2]. Die für die Zwecke dieser Stellungnahme verwendete Terminologie ist nachstehend in Abschnitt 2.2 und 2.3 angegeben.

4. Die von der irischen Aufsichtsbehörde (IE-AB) gestellten Fragen lauteten:

Frage 1: Ist von einem endgültigen KI-Modell, das unter Verwendung personenbezogener Daten trainiert wurde, in allen Fällen anzunehmen, dass es nicht unter die Definition des Begriffs „personenbezogene Daten“ (im Sinne von Artikel 4 Nummer 1 DSGVO) fällt?

Falls Frage 1 zu bejahen ist:

i. Ab welcher Phase der Verarbeitungsvorgänge, die zu einem KI-Modell führen, ist keine Verarbeitung personenbezogener Daten mehr gegeben?
   a) Wie kann Nachweis dafür erbracht werden, dass das KI-Modell keine personenbezogenen Daten verarbeitet?

ii. Gibt es Faktoren, die bewirken würden, dass der Betrieb des endgültigen KI-Modells nicht mehr als anonym anzusehen wäre?
   a) Falls ja: Wie kann für die im Hinblick auf diese Faktoren ergriffenen Minderungs-, Präventions- oder Schutzmaßnahmen (die sicherstellen sollen, dass das KI-Modell keine personenbezogenen Daten verarbeitet) Nachweis erbracht werden?

Falls Frage 1 zu verneinen ist:

i. Unter welchen Umständen könnte dies der Fall sein?
   a) Falls dem so ist: Wie kann Nachweis für die Maßnahmen erbracht werden, die ergriffen wurden, um sicherzustellen, dass das KI-Modell keine personenbezogenen Daten verarbeitet?

Frage 2: Wie sollte ein Verantwortlicher, der für die Verarbeitung personenbezogener Daten zur Erstellung, Aktualisierung und/oder Entwicklung eines KI-Modells die Rechtsgrundlage des berechtigten Interesses anführt, Nachweis dafür erbringen, dass das berechtigte Interesse eine geeignete Rechtsgrundlage bietet, und zwar sowohl in Bezug auf die Verarbeitung von Dritten erhobener Daten als auch in Bezug auf die Verarbeitung selbst erhobener Daten?

i. Welche Erwägungen sollte der Verantwortliche berücksichtigen, um sicherzustellen, dass die Interessen der betroffenen Personen, deren personenbezogene Daten verarbeitet werden, ordnungsgemäß gegen die Interessen des Verantwortlichen abgewogen werden, und zwar im Zusammenhang mit:
   a) von Dritten erhobenen Daten?

   b) von dem Verantwortlichen selbst erhobenen Daten?

[2] Ersuchen, S. 1.

Frage 3: Wie sollte, wenn personenbezogene Daten nach der Trainingsphase innerhalb eines KI-Modells oder eines KI-Systems, von dem das KI-Modell einen Teil darstellt, verarbeitet werden, ein Verantwortlicher, der die Rechtsgrundlage des berechtigten Interesses anführt, Nachweis dafür erbringen, dass die berechtigten Interessen eine geeignete Rechtsgrundlage bieten?

Frage 4: Falls festgestellt werden sollte, dass ein KI-Modell unter Verwendung rechtswidrig verarbeiteter personenbezogener Daten erstellt, aktualisiert oder entwickelt wurde: Hätte dies, wenn die Verarbeitungstätigkeit oder der Betrieb des KI-Modells, sei es allein oder als Teil eines KI-Systems, fortgesetzt wird oder später erfolgt, Folgen für deren Rechtmäßigkeit, und welche Folgen wären dies, wenn:

i. das KI-Modell, sei es allein oder als Teil eines KI-Systems, personenbezogene Daten verarbeitet?
ii. weder das KI-Modell noch das KI-Modell als Teil eines KI-Systems personenbezogene Daten verarbeitet?

1.2 Zulässigkeit des Ersuchens um Stellungnahme nach Artikel 64 Absatz 2 DSGVO

5. Gemäß Artikel 64 Absatz 2 DSGVO kann jede Aufsichtsbehörde beantragen, dass eine Angelegenheit mit allgemeiner Geltung oder mit Auswirkungen in mehr als einem Mitgliedstaat vom Ausschuss geprüft wird, um eine Stellungnahme zu erhalten.

6. Die ersuchende Aufsichtsbehörde stellte dem EDSA Fragen zu Datenschutzaspekten im Zusammenhang mit KI-Modellen. In ihrem Ersuchen wies sie darauf hin, dass zwar viele Organisationen inzwischen KI-Modelle, einschließlich großer Sprachmodelle („**LLM**"), in ihrem Geschäftsbetrieb verwenden, dass jedoch deren Betrieb, Training und Verwendung „einige weitreichende Datenschutzbedenken"[3] aufwürfen, die „Auswirkungen auf betroffene Personen in allen Teilen der EU und des EWR haben"[4].

7. Das Ersuchen wirft im Wesentlichen Fragen auf in Bezug auf: (i) die Anwendung des Begriffs „personenbezogene Daten"; (ii) den Grundsatz der Rechtmäßigkeit, insbesondere im Hinblick auf die Rechtsgrundlage des berechtigten Interesses im Zusammenhang mit KI-Modellen; sowie (iii) die Folgen einer rechtswidrigen Verarbeitung personenbezogener Daten in der Entwicklungsphase eines KI-Modells für die spätere Verarbeitungstätigkeit oder den späteren Betrieb des Modells.

8. Der Ausschuss ist der Ansicht, dass dieses Ersuchen eine „*Angelegenheit mit allgemeiner Geltung*" im Sinne von Artikel 64 Absatz 2 DSGVO betrifft. Insbesondere betrifft diese Sache die Auslegung und Anwendung von Artikel 4 Absatz 1, Artikel 5 Absatz 1 Buchstabe a und Artikel 6 DSGVO auf die Verarbeitung personenbezogener Daten in der Entwicklung und im Einsatz von KI-Modellen. Die Anwendung dieser Bestimmung auf KI-Modelle wirft nach Ansicht der ersuchenden Aufsichtsbehörde systemische, abstrakte und neuartige Fragen auf[5]. Die rapide Entwicklung und der Einsatz von KI-Modellen durch immer mehr Organisationen wirft spezifische Fragen auf, wobei, wie es im Ersuchen heißt, „*es für den EDSA sehr nützlich sein wird, wenn zu den mit diesem Ersuchen aufgeworfenen Fragen ein gemeinsamer Standpunkt erzielt wird, da diese Fragen für die kurz- und mittelfristig geplante Arbeit des EDSA von zentraler Bedeutung sind*"[6]. Zudem können KI-Technologien in einem breiten Spektrum

[3] Ersuchen, S. 1.
[4] Ebenda.
[5] Ersuchen, S. 2.
[6] Ersuchen, S. 1. Im Arbeitsprogramm des EDSA für 2024-2025, das am 8. Oktober 2024 beschlossen wurde und unter https://www.edpb.europa.eu/system/files/2024-10/edpb_work_programme_2024-2025_en.pdf

wirtschaftlicher und gesellschaftlicher Tätigkeiten vielfältige Chancen eröffnen und Nutzen bieten. Außerdem bietet die DSGVO einen Rechtsrahmen, der verantwortungsvolle Innovation fördert. Daraus folgt, dass es im allgemeinen Interesse liegt, diese Bewertung in Form einer Stellungnahme des EDSA vorzunehmen, um sicherzustellen, dass gewisse Bestimmungen der DSGVO im Zusammenhang mit KI-Modellen einheitliche Anwendung finden.

9. Die in Artikel 64 Absatz 2 DSGVO genannte alternative Voraussetzung betrifft Angelegenheiten *„mit Auswirkungen in mehr als einem Mitgliedstaat“*. Der EDSA erinnert daran, dass der Begriff „Auswirkungen“ weit auszulegen ist und deshalb nicht einfach auf die rechtlichen Auswirkungen beschränkt ist[7]. Da immer mehr KI-Modelle trainiert und von immer mehr Organisationen im EWR verwendet werden, hat dies Auswirkungen auf eine große Zahl betroffener Personen im gesamten EWR, von denen einige bereits ihrer zuständigen Aufsichtsbehörde gegenüber Bedenken angemeldet haben[8] Der EDSA ist deshalb der Ansicht, dass die von der ersuchenden Aufsichtsbehörde vorgelegte Angelegenheit auch diese Voraussetzung erfüllt.

10. Das Ersuchen enthält schriftliche Ausführungen über den Hintergrund (einschließlich des einschlägigen Rechtsrahmens) und die Begründung der dem Ausschuss vorgelegten Fragen. Der Ausschuss stellt fest, dass es sich folglich um ein begründetes Ersuchen im Sinne von Artikel 10 Absatz 3 der Geschäftsordnung des EDSA handelt.

11. Gemäß Artikel 64 Absatz 3 DSGVO[9] gibt der EDSA keine Stellungnahme ab, wenn er bereits eine Stellungnahme zu derselben Angelegenheit abgegeben hat. Der EDSA hat noch keine Stellungnahme zu derselben Angelegenheit abgegeben und hat die im Ersuchen gestellten Fragen noch nicht beantwortet.

12. Der Ausschuss hält das Ersuchen deshalb für zulässig und ist der Auffassung, dass die mit dem Ersuchen gestellten Fragen in dieser gemäß Artikel 64 Absatz 2 DSGVO angenommenen Stellungnahme („**Stellungnahme**“) untersucht werden sollten.

2 Anwendungsbereich und Schlüsselbegriffe

2.1 Anwendungsbereich der Stellungnahme

13. Der Ausschuss ist sich mit der ersuchenden Aufsichtsbehörde einig, dass Entwicklung und Einsatz von KI-Modellen in datenschutzrechtlicher Hinsicht grundlegende Fragen aufwerfen. Dabei geht es insbesondere darum: (i) wann und auf welche Weise ein KI-Modell für „anonym“ befunden werden kann (Frage 1 des Ersuchens); (ii) auf welche Weise Verantwortliche die Angemessenheit des berechtigten Interesses als Rechtsgrundlage nachweisen können, und zwar zum einen in der Entwicklungsphase (Frage 2 des Ersuchens) und zum anderen in der Einsatzphase (Frage 3 des Ersuchens); und (iii) ob eine rechtswidrige Verarbeitung personenbezogener Daten in der

abrufbar ist, ist vorgesehen, dass der EDSA unter anderem Leitlinien zur Anonymisierung, Pseudonymisierung und zum Data Scraping im Zusammenhang mit generativer KI herausgibt.

[7] EDPB, Internal document 3/2019 on Internal guidance on Article 64 (2) GDPR (Internes Dokument 3/2019 des EDSA zu internen Richtlinien zu Artikel 64 Absatz 2 DSGVO), angenommen am 8. Oktober 2019, Randnummer 15, abrufbar unter https://www.edpb.europa.eu/system/files/2022-07/internaledpb_document_201903_art64.2_en.pdf.

[8] Ersuchen, S. 1-2.

[9] Artikel 64 Absatz 3 DSGVO und Artikel 10 Absatz 4 der Geschäftsordnung des EDSA.

Entwicklungsphase Folgen für die Rechtmäßigkeit der späteren Verarbeitungstätigkeit oder des späteren Betriebs des KI-Modells hat (Frage 4 des Ersuchens).

14. Der EDSA erinnert daran, dass die Zuständigkeit für die Überwachung der DSGVO-Anwendung bei den Aufsichtsbehörden liegt und dass diese zur unionsweit einheitlichen Anwendung der DSGVO beitragen[10]. Es liegt deshalb im Zuständigkeitsbereich der Aufsichtsbehörden, spezifische KI-Modelle zu untersuchen und dabei Einzelfallprüfungen vorzunehmen.

15. Diese Stellungnahme setzt zuständigen Aufsichtsbehörden einen Rahmen für die Prüfung spezifischer Fälle, in denen sich die (oder einige der) mit dem Ersuchen aufgeworfenen Fragen stellen. Diese Stellungnahme soll nicht erschöpfend sein, sondern allgemeine Erwägungen zur Auslegung einschlägiger Bestimmungen bieten, denen die zuständigen Aufsichtsbehörden bei der Wahrnehmung ihrer Untersuchungsbefugnissen höchste Beachtung schenken sollten. Diese Stellungnahme ist an die zuständigen Aufsichtsbehörden gerichtet und betrifft deren Tätigkeiten und Befugnisse; die sich aus der DSGVO ergebenden Verpflichtungen der Verantwortlichen und Auftragsverarbeiter bleiben unberührt. Insbesondere sind die Verantwortlichen nach dem in Artikel 5 Absatz 2 DSGVO verankerten Grundsatz der Rechenschaftspflicht verpflichtet, sämtliche Grundsätze, die ihre Verarbeitung personenbezogener Daten betreffen, einzuhalten und für deren Einhaltung Nachweis erbringen zu können.

16. Auch wenn in dieser Stellungnahme in einigen Fällen Beispiele angeführt werden, wird es, weil die im Ersuchen gestellten Fragen sehr weit gefasst sind und verschiedene Arten von KI-Modellen darunter fallen, nicht möglich sein, alle in Betracht kommenden Szenarien in dieser Stellungnahme zu berücksichtigen. Bei der Auslegung der in dieser Stellungnahme gemachten Ausführungen des EDSA sollte berücksichtigt werden, dass sich die mit KI-Modellen verbundenen Technologien rapide weiterentwickeln.

17. **Nicht in dieser Stellungnahme berücksichtigt sind die nachstehenden Bestimmungen, die für die Bewertung der für KI-Modelle geltenden Datenschutzanforderungen ebenfalls eine wichtige Rolle spielen könnten:**

- **Verarbeitung besonderer Datenkategorien** Der EDSA erinnert an das in Artikel 9 Absatz 1 DSGVO geregelte Verbot der Verarbeitung besonderer Kategorien personenbezogener Daten und die in Artikel 9 Absatz 2 DSGVO geregelten engen Ausnahmetatbestände[11]. Dazu hat der Gerichtshof der Europäischen Union („**EuGH**“) präzisiert, dass *„[w]enn ein Datensatz, der sowohl sensible als auch nicht sensible Daten enthält, … als Ganzes erhoben wird, ohne dass die Daten zum Zeitpunkt dieser Erhebung voneinander getrennt werden können, … die Verarbeitung dieses Datensatzes aber als im Sinne von Art. 9 Abs. 1 DSGVO untersagt anzusehen [ist], sofern sie mindestens ein sensibles Datum umfasst und keine der in Art. 9 Abs. 2 DSGVO genannten Ausnahmen greift*“[12]. Des Weiteren hat der EuGH auch hervorgehoben, dass *„für die Zwecke der Anwendung der in Art. 9 Abs. 2 Buchst. e DSGVO vorgesehenen Ausnahme zu prüfen [ist], ob die*

[10] Artikel 51 Absatz 1 DSGVO und Artikel 51 Absatz 2 DSGVO.

[11] Vgl. auch den EDPB Report of the work undertaken by the ChatGPT Taskforce (Bericht des EDSA über die von der ChatGPT Taskforce unternommene Arbeit), angenommen am 23. Mai 2024, Randnummer 18: „Im Hinblick auf die Verarbeitung besonderer Kategorien personenbezogener Daten ist die Verarbeitung nur rechtmäßig, wenn zusätzlich einer der in Artikel 9 Absatz 2 genannten Ausnahmetatbestände erfüllt ist. *Einer dieser grundsätzlich in Betracht kommenden Ausnahmetatbestände ist Artikel 9 Absatz 2 Buchstabe e DSGVO. Die bloße Tatsache, dass personenbezogene Daten öffentlich zugänglich sind, bedeutet noch nicht, dass ‚die betroffene Person diese Daten offensichtlich öffentlich gemacht hat' …*“

[12] Urteil des EuGH vom 4. Juli 2023, Rechtssache C-252/21, *Meta/Bundeskartellamt* (ECLI:EU:C:2023:537), Randnummer 89.

betroffene Person die Absicht hatte, die fraglichen personenbezogenen Daten ausdrücklich und durch eine eindeutige bestätigende Handlung der breiten Öffentlichkeit zugänglich zu machen" [13] . Diese Erwägungen sind zu berücksichtigen, wenn die Verarbeitung personenbezogener Daten im Zusammenhang mit KI-Modellen besondere Datenkategorien betrifft.

- **Automatisierte Entscheidungsfindung einschließlich Profiling**: Die im Zusammenhang mit KI-Modellen durchgeführten Verarbeitungsvorgänge fallen unter Umständen in den Anwendungsbereich von Artikel 22 DSGVO, der den Verantwortlichen zusätzliche Pflichten auferlegt und den betroffenen Personen zusätzliche Garantien bietet. Insoweit erinnert der EDSA an seine Leitlinien zu automatisierten Entscheidungen im Einzelfall einschließlich Profiling für die Zwecke der Verordnung 2016/679[14].

- **Vereinbarkeit der Zwecke**: Für bestimmte Rechtsgrundlagen sieht Artikel 6 Absatz 4 DSGVO die Kriterien vor, die ein Verantwortlicher berücksichtigen muss, um festzustellen, ob die Verarbeitung zu einem anderen Zweck mit dem Zweck, zu dem die personenbezogenen Daten ursprünglich erhoben wurden, vereinbar ist. Diese Bestimmung kann im Zusammenhang mit der Entwicklung und dem Einsatz von KI-Modellen relevant sein, und ihre Anwendbarkeit sollte von den Aufsichtsbehörden geprüft werden.

- **Datenschutz-Folgenabschätzungen** („**DSFA**") (Artikel 35 DSGVO): Soweit die Verarbeitung im Zusammenhang mit KI-Modellen wahrscheinlich zu einem hohen Risiko für die Rechte und Freiheiten natürlicher Personen führt, sind DSFA ein wichtiges Element der Rechenschaftspflicht[15].

- **Grundsatz des Datenschutzes durch Technikgestaltung** (Artikel 25 Absatz 1 DSGVO): Datenschutz durch Technikgestaltung ist eine wesentliche Garantie, die von den Aufsichtsbehörden im Zusammenhang mit der Entwicklung und dem Einsatz von KI-Modellen zu prüfen ist.

2.2 Schlüsselbegriffe

18. Vorab möchte der EDSA – ausschließlich für die Zwecke dieser Stellungnahme – die von ihm in dieser Stellungnahme verwendeten Begriffe und Konzepte klären:

- „**Selbst erhobene Daten**" bezeichnet personenbezogene Daten, die der Verantwortliche selbst von den betroffenen Personen erhoben hat.

- „**Von Dritten erhobene Daten**" bezeichnet personenbezogene Daten, die der Verantwortliche nicht selbst von den betroffenen Personen erlangt hat, sondern die er von einem Dritten (z. B. einem Datenmakler) erhoben oder empfangen hat oder die durch Web Scraping erhoben wurden.

[13] Urteil des EuGH vom 4. Juli 2023, Rechtssache C-252/21, *Meta/Bundeskartellamt* (ECLI:EU:C:2023:537), Randnummer 77.

[14] Artikel-29-Datenschutzgruppe („**WP 29**"), Leitlinien zu automatisierten Entscheidungen im Einzelfall einschließlich Profiling für die Zwecke der Verordnung 2016/679, zuletzt überarbeitet und angenommen am 6. Februar 2018, vom EDSA gebilligt am 25. Mai 2018. Vgl. auch Urteil des EuGH vom 7. Dezember 2023, Rechtssache C-634/21, SCHUFA Holding und andere (ECLI:EU:C:2023:957).

[15] Siehe WP29, Leitlinien zur Datenschutz-Folgenabschätzung (DSFA) und Beantwortung der Frage, ob eine Verarbeitung im Sinne der Verordnung 2016/679 „wahrscheinlich ein hohes Risiko mit sich bringt", zuletzt überarbeitet und angenommen am 4. Oktober 2017, vom EDSA gebilligt am 25. Mai 2018.

- „**Web Scraping**“ ist eine allgemein gebräuchliche Technik für die Datenextraktion aus öffentlich verfügbaren Online-Quellen. Informationen, die beispielsweise von Diensten wie Nachrichtenmedien, sozialen Medien, Diskussionsforen und persönlichen Websites ausgelesen werden, können personenbezogene Daten enthalten.

- Das Ersuchen betrifft den **„Lebenszyklus“ von KI-Modellen** sowie verschiedene Phasen des Lebenszyklus, unter anderem „Erstellung“, „Entwicklung“, „Training“, „Aktualisierung“, „Fine-Tuning“, „Betrieb“ oder „Post-Training“ von KI-Modellen. Der EDSA erkennt an, dass diese Phasen gegebenenfalls in die Entwicklungsphase und in die Einsatzphase von KI-Modellen fallen können und dass dabei möglicherweise personenbezogene Daten zu den verschiedensten Verarbeitungszwecken verarbeitet werden. Für die Zwecke dieser Stellungnahme hält es der EDSA jedoch für wichtig, die wahrscheinlich vorkommenden Phasen straffer zu kategorisieren. Deshalb verwendet der EDSA für die Zwecke dieser Stellungnahme die Kategorien „**Entwicklungsphase**“ und „**Einsatzphase**“. Die Entwicklung eines KI-Modells umfasst sämtliche Phasen vor dem Einsatz des KI-Modells; darunter fallen unter anderem die Codeentwicklung, die Erhebung personenbezogener Trainingsdaten, die Vorverarbeitung personenbezogener Trainingsdaten und das Training. Der Einsatz eines KI-Modells umfasst sämtliche Phasen im Zusammenhang mit der Verwendung des KI-Modells und kann alle nach der Entwicklungsphase ausgeführten Vorgänge beinhalten. Der EDSA bleibt sich der Vielfalt der Anwendungsfälle und deren potenziellen Auswirkungen auf die Verarbeitung personenbezogener Daten bewusst; die Aufsichtsbehörden sollten deshalb prüfen, ob die Ausführungen in dieser Stellungnahme für die Verarbeitung, mit der sie gerade befasst sind, relevant sind.

- Der EDSA hebt auch hervor, dass der Begriff „**Training**“ erforderlichenfalls den Teil der Entwicklungsphase betrifft, in dem KI-Modelle aus Daten lernen, um die für sie vorgesehene Aufgabe zu erfüllen (was im nächsten Abschnitt dieser Stellungnahme erklärt wird).

- Der Begriff „**KI-Modell**“ und sein Anwendungsbereich, so wie ihn der EDSA für die Zwecke dieser Stellungnahme versteht, wird im folgenden Abschnitt näher erklärt.

2.3 Verwendung des Begriffs „KI-Modelle“ in dieser Stellungnahme

19. In der EU-Verordnung über künstliche Intelligenz („**KI-Verordnung**“)[16] ist der Begriff „KI-System“ definiert als „*ein maschinengestütztes System, das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist und das nach seiner Betriebsaufnahme anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können*“[17]. Weitere Erklärungen zum Begriff „KI-System“ sind in Erwägungsgrund 12 der KI-Verordnung zu finden. Ein wesentliches Merkmal von KI-Systemen ist also ihre Ableitungsfähigkeit. Zu den Techniken, die während der Gestaltung eines KI-Systems das Ableiten ermöglichen, gehören Ansätze für maschinelles Lernen sowie logik- und wissensgestützte Konzepte.

20. Für den Begriff „KI-Modell“ bietet die KI-Verordnung jedoch nur eine indirekte Definition: „*Obwohl KI-Modelle wesentliche Komponenten von KI-Systemen sind, stellen sie für sich genommen keine KI-Systeme dar. Damit KI-Modelle zu KI-Systemen werden, ist die Hinzufügung weiterer Komponenten,*

[16] Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz).
[17] Artikel 3 Absatz 1 der KI-Verordnung.

zum Beispiel einer Nutzerschnittstelle, erforderlich. KI-Modelle sind in der Regel in KI-Systeme integriert und Teil davon".[18]

21. Nach dem Verständnis des EDSA ist der Begriff „KI-Modell" im Ersuchen enger definiert als in der KI-Verordnung; im Ersuchen heißt es zum Begriff „KI-Modell", dass dieser „*das Produkt umfasst, das sich aus den Trainingsmechanismen ergibt, die im Zusammenhang mit künstlicher Intelligenz, maschinellem Lernen, tiefem Lernen (Deep Learning) oder anderen damit verbundenen Verarbeitungskontexten auf einen Trainingsdatensatz angewendet werden*", wobei präzisiert wird, dass „*[d]er Begriff … sowohl für KI-Modelle [gilt], die dafür bestimmt sind, weiterem Training oder Feintuning unterzogen und/oder weiterentwickelt zu werden, als auch für KI-Modelle, für die dies nicht der Fall ist.*"[19]

22. Auf dieser Grundlage ist der EDSA bei der Annahme dieser Stellungnahme davon ausgegangen, dass ein KI-System, um das ihm gesetzte Ziel zu erreichen, auf ein KI-Modell angewiesen ist, das in das größere Rahmenwerk des KI-Systems integriert ist (ein Beispiel dafür wäre ein KI-System für den Kundendienst, bei dem für die Beantwortung von Nutzeranfragen ein mit historischen Gesprächsdaten trainiertes KI-Modell verwendet wird).

23. Des Weiteren sind die für diese Stellungnahme relevanten KI-Modelle (im Folgenden auch „**Modelle**" genannt) solche, die durch ein Trainingsverfahren entwickelt werden. Ein solches Trainingsverfahren ist Teil der Entwicklungsphase, in der die Modelle aus Daten lernen, um die für sie vorgesehene Aufgabe zu erfüllen. Für das Trainingsverfahren ist ein Datensatz erforderlich, anhand dessen das Modell Muster erkennt und „lernt". In diesen Fällen wird das Modell verschiedene Techniken verwenden, um eine Darstellung des aus dem Trainingsdatensatz entnommenen Wissens aufzubauen. Dies gilt insbesondere für das maschinelle Lernen.

24. Praktisch betrachtet ist jedes KI-Modell ein Algorithmus, dessen Funktionsweise durch mehrere Elemente bestimmt wird. Beispielsweise handelt es sich bei Modellen für tiefes Lernen (Deep-Learning-Modelle) der Form nach häufig um ein neuronales Netz mit mehreren Rechenschichten, die aus Knoten bestehen, die über gewichtete Kanten verbunden sind, wobei die Gewichtungen im Zuge des Trainings angepasst werden, wodurch Beziehungen zwischen den Eingaben und Ausgaben erlernt werden. Die Eigenschaften eines einfachen Deep-Learning-Modells wären also: (i) Art und Größe jeder Rechenschicht, (ii) die jeder Kante zugewiesene Gewichtung (auch als „Parameter" bezeichnet), (iii) die Aktivierungsfunktionen[20] zwischen Rechenschichten sowie möglicherweise (iv) andere zwischen den Rechenschichten stattfindende Operationen. Wird beispielsweise ein einfaches Deep-Learning-Modell für die Bildklassifizierung trainiert, so werden die Eingaben (die „**Bild-Pixel**") mit den Ausgaben in Beziehung gesetzt; die Gewichtungen können angepasst werden und die Anpassung soll dafür sorgen, dass die Ausgabe in den meisten Fällen richtig ist.

25. Andere Beispiele für Deep-Learning-Modelle sind große Sprachmodelle (Large Language Models, LLM) und generative KI, die z. B. dazu verwendet werden, Inhalte zu generieren, die von Menschen produzierten Inhalten vergleichbar sind, und neue Daten zu schaffen.

26. **Im Hinblick auf die vorgenannten Erwägungen und im Einklang mit dem Ersuchen beschränkt sich der Anwendungsbereich dieser Stellungnahme auf die Untergruppe derjenigen KI-Modelle, die für ihre Erstellung mit personenbezogenen Daten trainiert wurden.**

[18] Siehe Erwägungsgrund 97 der KI-Verordnung.
[19] Ersuchen, S. 3.
[20] Dies sind Funktionen, die auf Grundlage der Eingaben und Gewichtungen den Output eines neuronalen Knotens berechnen, der dann an die nächste Rechenschicht des neuronalen Netzes gesendet wird.

3 Begründetheit des Ersuchens

3.1 KI-Modelle im Verhältnis zur Definition des Begriffs „personenbezogene Daten“

27. In Artikel 4 Ziffer 1 DSGVO sind personenbezogene Daten definiert als *„alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person (im Folgenden ‚betroffene Person‘) beziehen“.* In Erwägungsgrund 26 DSGVO heißt es, dass die Datenschutzgrundsätze nicht auf anonyme Daten – d. h. auf Informationen, die sich nicht auf eine identifizierte oder identifizierbare natürliche Person beziehen, – anwendbar sind; dabei werden hinsichtlich der Identifizierbarkeit *„alle Mittel berücksichtigt …, die von dem Verantwortlichen oder einer anderen Person nach allgemeinem Ermessen wahrscheinlich genutzt werden*“. Dazu gehören: (i) Daten, die niemals auf eine identifizierte oder identifizierbare natürliche Person bezogen waren, sowie (ii) personenbezogene Daten, die auf solche Weise anonymisiert wurden, dass die betroffene Person nicht oder nicht mehr identifizierbar ist.

28. Frage 1[21] des Ersuchens kann also beantwortet werden, indem man prüft, ob ein KI-Modell, das durch ein Training erstellt wurde, bei dem personenbezogene Daten verarbeitet wurden, in allen Fällen als anonym anzusehen sein sollte. Im Hinblick auf den Wortlaut der Frage wird der EDSA in diesem Abschnitt vom Prozess für das „Training“ eines KI-Modells sprechen.

29. Zunächst möchte der EDSA vor allem die folgenden allgemeinen Erwägungen voranschicken. KI-Modelle sind – unabhängig davon, ob sie mit personenbezogenen Daten trainiert werden oder nicht – in der Regel darauf ausgelegt, Vorhersagen zu machen oder Schlüsse zu ziehen, d. h. sie sind darauf ausgelegt, abzuleiten. Des Weiteren sind KI-Modelle, die mit personenbezogenen Daten trainiert werden, oftmals darauf ausgelegt, Schlüsse im Hinblick auf natürliche Personen zu ziehen, wobei dies andere Personen sind als diejenigen, deren personenbezogene Daten für das Training des KI-Modells verwendet wurden. Einige KI-Modelle sind allerdings eigens darauf ausgelegt, personenbezogene Daten über diejenigen natürlichen Personen, mit deren personenbezogenen Daten das KI-Modell trainiert wurde, auszugeben oder derartige Daten in irgendeiner Weise bereitzustellen. In solchen Fällen ist es eine den KI-Modellen (in der Regel zwangsläufig) innewohnende Eigenschaft, dass sie Daten über eine identifizierte oder identifizierbare natürliche Person enthalten; folglich wird es sich um die Verarbeitung personenbezogener Daten handeln. Solche Arten von KI-Modellen können nicht als anonym angesehen werden. Beispiele dafür wären etwa (i) ein generatives Modell, bei dem für das Fine-Tuning Stimmaufzeichnungen einer Person verwendet werden, um deren Stimme nachzuahmen; oder (ii) Modelle, die darauf ausgelegt sind, als Antwort auf Fragen zu einer bestimmten Person personenbezogene Daten aus dem Training auszugeben.

30. Im Hinblick auf die vorstehenden Erwägungen fokussiert der EDSA in seiner Antwort auf die Frage 1 des Ersuchens auf die Situation der KI-Modelle, die nicht darauf ausgelegt sind, im Zusammenhang mit den Trainingsdaten personenbezogene Daten auszugeben.

31. Nach Ansicht des EDSA kann es, selbst wenn ein KI-Modell nicht absichtlich darauf ausgelegt wurde, Informationen aus den Trainingsdaten auszugeben, die sich auf eine identifizierte oder identifizierbare natürliche Person beziehen, doch vorkommen, dass Daten aus dem Trainingsdatensatz – einschließlich personenbezogener Daten – nach wie vor in den Parametern des Modells „absorbiert“ (d. h. durch mathematische Objekte repräsentiert) sind. Es mag sein, dass diese Daten sich von den ursprünglichen

[21] *„Ist von einem endgültigen KI-Modell, das unter Verwendung personenbezogener Daten trainiert wurde, in allen Fällen anzunehmen, dass es nicht unter die Definition des Begriffs „personenbezogene Daten“ (im Sinne von Artikel 4 Nummer 1 DSGVO) fällt?*“

Trainingsdatenpunkten unterscheiden, aber dennoch weiterhin die ursprünglichen Informationen dieser Daten enthalten, die dann möglicherweise letztendlich direkt oder indirekt aus dem Modell entnommen oder in sonstiger Weise daraus erlangt werden können. Soweit es möglich ist, mit Mitteln, die nach allgemeinem Ermessen wahrscheinlich genutzt werden, aus einem KI-Modell Informationen zu erlangen, die sich auf eine identifizierte oder identifizierbare natürliche Person beziehen, deren personenbezogene Daten für das Training eines Modells verwendet wurden, lässt dies den Schluss zu, dass ein derartiges Modell nicht anonym ist.

32. Dazu heißt es im Ersuchen, dass *„[I]n veröffentlichten Forschungsarbeiten auf einige potenzielle Schwachstellen hingewiesen wird, die bei KI-Modellen auftreten und dazu führen könnten, dass personenbezogene Daten verarbeitet werden*[22] *bzw. die Verarbeitung personenbezogener Daten auch noch weitergeht, wenn die Modelle für die Verwendung mit anderen Daten eingesetzt werden, sei es durch Anwendungsprogrammierschnittstellen (‚**API**‘) oder durch Schnittstellen für die Eingabeaufforderung (‚Prompts‘)“*[23].

33. Auch die Forschung zur Trainingsdatenextraktion ist sehr dynamisch[24]. Sie zeigt, dass es in einigen Fällen mit Mitteln, die aller Wahrscheinlichkeit nach genutzt werden, bei einigen KI-Modellen möglich ist, personenbezogene Daten zu entnehmen oder aber einfach im Zuge der Interaktion mit einem KI-Modell (das beispielsweise Teil eines KI-Systems ist) versehentlich personenbezogene Daten zu erlangen. Die weitere Forschung auf diesem Gebiet wird eine genauere Bewertung der im Einzelfall bestehenden Restrisiken der Regurgitation [25] (Auswerfen von Trainingsdaten) und Extraktion personenbezogener Daten erleichtern.

34. **Im Hinblick auf die vorstehenden Erwägungen ist der EDSA der Ansicht, dass mit personenbezogenen Daten trainierte KI-Modelle nicht in allen Fällen für anonym befunden werden können. Vielmehr ist die Bestimmung, ob ein KI-Modell anonym ist, jeweils im Einzelfall auf Grundlage spezifischer Kriterien vorzunehmen.**

3.2 Umstände, bei deren Vorliegen KI-Modelle für anonym befunden werden könnten, und der dafür zu erbringende Nachweis

[22] Wie z. B. Membership Inference Attacks (um herauszufinden, ob ein bestimmter Datenpunkt Teil des zur Erzeugung eines Modells verwendeten Trainingsdatensatzes war) (OWASP), und Model Inversion Attacks (um synthetische Datenpunkte oder Klassenrepräsentationen zu erzeugen) (OWASP u. Veale et al, 2018).

[23] Ersuchen, S. 1-2.

[24] Vgl. dazu beispielsweise: (i) Veale M., Binns R., Edwards L., 2018, *Algorithms that remember: model inversion attacks and data protection law*. Phil. Trans. R. Soc. A 376: 20180083, abrufbar unter http://dx.doi.org/10.1098/rsta.2018.0083; (ii) Brown H., Lee K., Mireshghallah F., Shokri R., und Tramèr F., *What Does it Mean for a Language Model to Preserve Privacy?*, 2022, ACM Digital Library, FAccT ’22, 20. Juni 2022, Seoul, Republik Korea, abrufbar unter https://dl.acm.org/doi/abs/10.1145/3531146.3534642; (iii) Vassilev A., Oprea A., Fordyce A., Anderson H., *Adversarial Machine Learning A Taxonomy and Terminology of Attacks and Mitigations*, Januar 2024, National Institute of Standards and Technology, abrufbar unter https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-2e2023.pdf; (iv) Carlini N., Tramèr F., Wallace E., Jagielski M., Herbert-Voss A., Lee K., Roberts A., Brown T., Song D., Erlingsson U., Oprea A., Raffel C., *Extracting Training Data from Large Language Models*, arXiv:2012.07805v2 (cs.CR) 15. Juni 2021, abrufbar unter https://arxiv.org/pdf/2012.07805; (v) Fredrikson M., Jha S., Ristenpart T., *Model Inversion Attacks that Exploit Confidence Information and Basic Countermeasures*, ACM Digital Library, 12. Oktober 2015, abrufbar unter https://dl.acm.org/doi/abs/10.1145/2810103.2813677; (vi) Zhang Y., Jia R., Pei H., Wang W., Li B., Song D., *The Secret Revealer: Generative Model-Inversion Attacks Against Deep Neural Networks*, arXiv:1911.07135v2 (cs.LG) 18. April 2020, abrufbar unter https://arxiv.org/pdf/1911.07135.

[25] Bei einem auf generativer KI basierenden KI-System spricht man von Regurgitation, wenn sich Ausgaben direkt auf Trainingsdaten beziehen.

35. Mit Frage 1 des Ersuchens[26] wird der EDSA ersucht, klarzustellen, unter welchen Umständen ein mit personenbezogenen Daten trainiertes KI-Modell für anonym befunden werden kann. Mit der Frage 1(i)(a) des Ersuchens [27] wird der EDSA ersucht, klarzustellen, welche Beweise und/oder Dokumentation die Aufsichtsbehörden bei der Prüfung, ob ein KI-Modell anonym ist, berücksichtigen sollten.

3.2.1 Allgemeine Erwägung hinsichtlich der Anonymisierung im vorliegenden Zusammenhang

36. Aus der Wendung „*alle Informationen*" in der Begriffsbestimmung für „*personenbezogene Daten*" in Artikel 4 Nummer 1 DSGVO ist ersichtlich, dass eine weite Auslegung dieses Begriffs angestrebt wird, die sämtliche Arten von Informationen beinhaltet, soweit sich diese auf die betroffene Person, die direkt oder indirekt identifiziert oder identifizierbar ist, „*beziehen*".

37. Die Informationen können sich auf eine natürliche Person beziehen, selbst wenn sie in solcher Weise technisch organisiert oder kodiert sind (beispielsweise in einem – sei es geschützten oder offenen – ausschließlich maschinenlesbaren Format), dass die Beziehung zu der betreffenden natürlichen Person nicht ohne Weiteres ersichtlich ist. In solchen Fällen kann es sein, dass Software-Anwendungen dazu verwendet werden, bestimmte Daten auf einfache Weise zu ermitteln, zu erkennen und zu entnehmen. Dies betrifft insbesondere KI-Modelle, bei denen Parameter statische Beziehungen zwischen den Trainingsdaten darstellen, wobei es möglich sein mag, exakte oder ungenaue (weil statistisch abgeleitete) personenbezogene Daten zu entnehmen, sei es direkt aus den Beziehungen zwischen den im Modell enthaltenen Daten oder durch Anfragen an das Modell.

38. Da KI-Modelle normalerweise keine Aufzeichnungen enthalten, die direkt isoliert oder verknüpft werden können, sondern Parameter, die probabilistische Beziehungen zwischen den im Modell enthaltenen Daten darstellen, gibt es realistische Szenarien, in denen es möglich sein mag, (z. B. durch Membership Inference) Informationen aus dem Modell abzuleiten[28]. Bevor sich eine Aufsichtsbehörde

[26] „*Unter welchen Umständen könnte dies der Fall sein?*"

[27] „*Falls dem so ist: Wie kann Nachweis für die Maßnahmen erbracht werden, die ergriffen wurden, um sicherzustellen, dass das KI-Modell keine personenbezogenen Daten verarbeitet?*"

[28] (i) Carlini N., Chien S., Nasr M., Song S., Terzis A., Tramer F., *Membership Inference Attacks From First Principles,* arXiv:2112.03570, abrufbar unter:https://arxiv.org/abs/2112.03570;
(ii) Creţu A. M., Guépin F., und De Montjoye Y. A., *Correlation inference attacks against machine learning models*. Sci. Adv.10, eadj9260(2024). DOI:10.1126/sciadv.adj9260 abrufbar unter https://www.science.org/doi/10.1126/sciadv.adj9260;
(iii) Dana L., Pydi M. S., Chevaleyre Y., *Memorization in Attention-only Transformers* arXiv:2411.10115v1 [cs.AI] 15. November 2024, abrufbar unter: https://arxiv.org/abs/2411.10115;
(iv) Gehrke M., Liebenow J., Mohammadi E., Braun T. et al. *Lifting in Support of Privacy-Preserving Probabilistic Inference*. Künstl Intell, 13. Juni 2024, abrufbar unter https://doi.org/10.1007/s13218-024-00851-y:
(v) Hu H., *Membership Inference Attacks and Defenses on Machine Learning Models Literature*, abrufbar unter: https://github.com/HongshengHu/membership-inference-machine-learning-literature;
(vi) Nasr M., Carlini N., Hayase J., Jagielski M., Cooper A. F., Ippolito D., Choquette-Choo C. A., Wallace E., Tramèr F. und Lee K., *Scalable Extraction of Training Data from (Production) Language Models*, arXiv:2311.17035 28. November 2023, abrufbar unter https://arxiv.org/abs/2311.17035;
(vii) Shokri R., Stronati M., Song C., Shmatikov V., *Membership Inference Attacks against Machine Learning Models* arXiv:1610.05820v2 [cs.CR], 31. März 2017, abrufbar unter https://arxiv.org/abs/1610.05820;
(viii) Staab R., Vero M., Mislav Balunović, Martin Vechev, 2024, *Beyond Memorization: Violating Privacy Via Inference with Large Language Models*, arXiv:2310.07298v2, 6. Mai 2024, abrufbar unter https://arxiv.org/abs/2310.07298;

der Auffassung des Verantwortlichen anschließt, dass dessen KI-Modell für anonym befunden werden kann, sollte sie deshalb zumindest prüfen, ob ihr hinreichend Nachweise dafür vorliegen, (i) dass es nicht möglich ist, personenbezogene Daten, die sich auf die Trainingsdaten beziehen, aus dem Modell zu entnehmen[29]; und (ii) dass die Ausgaben, die auf Anfragen an das Modell hin erzeugt werden, sich nicht auf die betroffenen Personen beziehen, mit deren personenbezogenen Daten das Modell trainiert wurde.

39. Bei der Prüfung, ob diese Voraussetzungen erfüllt sind, sollten die Aufsichtsbehörden drei Kriterien berücksichtigen:

40. Erstens sollten die Aufsichtsbehörden die in den jüngsten WP29-Stellungnahmen und/oder EDSA-Leitlinien genannten Kriterien prüfen. Was die Anonymisierung angeht, sollten die Aufsichtsbehörden nach dem Stand zum Zeitpunkt dieser Stellungnahme die Kriterien berücksichtigen, die in der WP29-Stellungnahme 5/2014 zu Anonymisierungstechniken („**WP29-Stellungnahme 5/2014**“) genannt sind. Danach können Daten für anonym befunden werden, sofern es nicht möglich ist, Informationen aus dem angeblich anonymen Datensatz herauszugreifen (zu isolieren), zu verknüpfen bzw. abzuleiten[30]. Des Weiteren wird dort ausgeführt, dass, wann immer *„ein Vorschlag ein Kriterium nicht [erfüllt], … eine gründliche Evaluierung der hinsichtlich einer Identifizierung bestehenden Risiken vorgenommen werden [sollte]“*[31]. **Wegen der bereits erwähnten Wahrscheinlichkeit der Extraktion und Ableitung hält es der EDSA für sehr wahrscheinlich, dass bei KI-Modellen eine sehr gründliche Prüfung der Identifizierungsrisiken erforderlich ist.**

41. Zweitens sollte die Prüfung alle Mittel berücksichtigen, die „vernünftigerweise entweder von dem Verantwortlichen für die Verarbeitung oder von einem ‚Dritten' eingesetzt werden könnten“[32], um natürliche Personen zu identifizieren, wobei für die Bestimmung dieser Mittel, wie in Erwägungsgrund 26 DSGVO erklärt wird, auf objektive Faktoren abgestellt werden sollte; objektive Faktoren sind beispielsweise:
    a. die Merkmale der eigentlichen Trainingsdaten, des KI-Modells und des Trainingsverfahrens[33];
    b. der Kontext, in dem das KI-Modell freigegeben und/oder verarbeitet wird[34];
    c. die zusätzlichen Informationen, anhand derer die Identifizierung möglich wäre und die der betreffenden Person zur Verfügung stehen;

(ix) Wu F., Cui L., Yao S., Yu S., *Inference Attacks in Machine Learning as a Service: A Taxonomy, Review, and Promising Directions* arXiv:2406.02027v1 [cs.LG], 27. Juni 2024, abrufbar unter https://arxiv.org/abs/2406.02027v1;
(x) Zhang J., Das D., Kamath G., Tramèr F., *Membership Inference Attacks Cannot Prove that a Model Was Trained On Your Data* arXiv:2409.19798v1, [cs.LG], 29. September 2024, abrufbar unter https://arxiv.org/abs/2409.19798;
(xi) Zhou Z., Xiang J., Chen C. und Su S., *Quantifying and Analyzing Entity-Level Memorization in Large Language Models*, arXiv:2308.15727v2 [cs.CL] 5. November 2023, abrufbar unter: https://arxiv.org/abs/2308.15727.

[29] Extraktion betrifft insbesondere den Fall, dass personenbezogene Daten aus dem KI-Modell selbst abgeleitet werden, wobei nur wenig oder gar kein Gebrauch von den Anfrageschnittstellen gemacht wird.
[30] WP29-Stellungnahme 05/2014, Seite 29.
[31] WP29-Stellungnahme 05/2014, Seite 29.
[32] Urteil des EuGH vom 19. Oktober 2016, Rechtssache C-582/14, *Breyer / Bundesrepublik Deutschland* (ECLI:EU:C:2016:779), Randnummer 43.
[33] Dazu zählen Merkmale wie die Einmaligkeit der Aufzeichnungen in den Trainingsdaten, die Genauigkeit der Informationen, die Aggregierung, die Randomisierung und insbesondere die Art und Weise, wie sich diese auf die Identifizierungsgefahr auswirken.
[34] Dazu gehören auch kontextbezogene Elemente, etwa die Beschränkung des Zugangs auf nur einige Personen sowie rechtliche Garantien.

d. der Kosten- und Zeitaufwand, der der Person entstünde, um solche zusätzlichen Informationen einzuholen (sofern ihr diese nicht bereits vorliegen)[35]; sowie
e. die zum Zeitpunkt der Verarbeitung verfügbare Technologie und technologische Entwicklungen[36].

42. Drittens sollten die Aufsichtsbehörden prüfen, ob die Verantwortlichen das Risiko der Identifizierung durch den Verantwortlichen und sonstige Arten „*anderer Personen*" (einschließlich des nicht unbeabsichtigten Zugangs Dritter zum KI-Modell) berücksichtigt haben, wobei auch zu prüfen ist, ob vernünftigerweise anzunehmen ist, dass diese in der Lage sind, sich Zugang zu den betreffenden Daten zu verschaffen oder dies betreffenden Daten zu verarbeiten.

43. **Insgesamt gelangt der EDSA zu dem Ergebnis, dass ein KI-Modell für anonym befunden werden kann, sofern** beim Einsatz von Mitteln, die nach allgemeinem Ermessen wahrscheinlich genutzt werden, **sowohl (i) die Wahrscheinlichkeit einer direkten (einschließlich einer probabilistischen) Entnahme personenbezogener Daten über die natürlichen Personen, deren personenbezogene Daten für das Training des Modells verwendet wurden, als auch (ii) die Wahrscheinlichkeit, dass derartige personenbezogene Daten durch Anfragen vorsätzlich oder nicht vorsätzlich erlangt werden, für jede betroffene Person vernachlässigbar gering[37] sein sollte. Die Aufsichtsbehörden sollten grundsätzlich davon ausgehen, dass bei KI-Modellen die Wahrscheinlichkeit der Identifizierung wahrscheinlich gründlich zu prüfen ist, um über die Frage ihrer etwaigen Anonymität entscheiden zu können. Die Wahrscheinlichkeitsprüfung sollte „alle Mittel berücksichtig[en], die von dem Verantwortlichen oder einer anderen Person nach allgemeinem Ermessen wahrscheinlich genutzt werden", wobei auch die unbeabsichtigte (Wieder-)Verwendung oder Offenlegung des Modells zu bedenken ist.**

3.2.2 Kriterien für die Bewertung der Restwahrscheinlichkeit der Identifizierung

44. Sowohl in der Entwicklungsphase als auch in der Einsatzphase eines KI-Modells können Maßnahmen ergriffen werden, um die Wahrscheinlichkeit, dass jemand daraus personenbezogene Daten erlangt, zu reduzieren; die Bewertung der Anonymität von KI-Modellen sollte jedoch auch die Möglichkeit des direkten Zugangs zum Modell bedenken.

45. Darüber hinaus sollten die Aufsichtsbehörden jeweils im Einzelfall bewerten, ob die Maßnahmen, mit denen der Verantwortliche sicherstellen und nachweisen will, dass das KI-Modell anonym ist, geeignet und wirksam sind.

46. Dabei kann das Ergebnis, zu dem die Aufsichtsbehörde gelangt, unterschiedlich ausfallen, je nachdem, ob es sich um ein für die Allgemeinheit verfügbares KI-Modell handelt, das für eine nicht bekannte Anzahl Menschen zugänglich ist, die über eine nicht bekannte Bandbreite von Methoden verfügen, mit denen sie versuchen könnten, personenbezogene Daten zu extrahieren, oder ob es sich um ein internes KI-Modell handelt, das nur Mitarbeitenden zugänglich ist. In beiden Fällen sollten die Aufsichtsbehörden überprüfen, dass die Verantwortlichen ihre Rechenschaftspflichten aus Artikel 5 Absatz 2 und Artikel 24 DSGVO erfüllt haben; allerdings könnten die „Mittel", die von anderen Personen „*nach allgemeinem Ermessen wahrscheinlich genutzt werden*" Auswirkungen darauf haben,

[35] Urteil des EuGH vom 7. März 2024, Rechtssache C-479/22 P, *OC/Kommission* (ECLI:EU:C:2024:215), Randnummer 50.
[36] Urteil des EuGH vom 7. März 2024, Rechtssache C-479/22 P, *OC/Kommission* (ECLI:EU:C:2024:215), Randnummer 50.
[37] Urteil des EuGH vom 19. Oktober 2016, Rechtssache C-582/14, *Breyer / Bundesrepublik Deutschland* (ECLI:EU:C:2016:779), Randnummer 46, und Urteil des EuGH vom 7. März 2024, Rechtssache C-479/22 P, *OC/Kommission* (ECLI:EU:C:2024:215), Randnummer 51.

welche Bandbreite und Art möglicher Szenarien zu betrachten sind. Je nach dem Kontext, in dem ein Modell entwickelt und eingesetzt wird, können die von den Aufsichtsbehörden angelegten Prüfungsmaßstäbe und die Anforderungen an die Angriffsresistenz unterschiedlich ausfallen.

47. Insoweit bietet der EDSA eine unverbindliche und nicht erschöpfende Liste der in Betracht kommenden Kriterien, die von den Aufsichtsbehörden für die Bewertung der von einem Verantwortlichen behaupteten Anonymität in Betracht gezogen werden können. Auch andere Ansätze können in Betracht kommen, sofern sie – insbesondere unter Berücksichtigung des Stands der Technik – ein gleichwertiges Schutzniveau bieten.

48. Für die Bewertung der Anonymität eines KI-Modells ist das Vorliegen oder Nichtvorliegen der nachstehend aufgeführten Elemente kein schlüssiges Kriterium.

3.2.2.1 Gestaltung des KI-Modells

49. Was die Gestaltung des KI-Modells angeht, sollten die Aufsichtsbehörden die Ansätze bewerten, denen die Verantwortlichen in der Entwicklungsphase gefolgt sind. Insoweit sollte die Anwendung und Wirksamkeit in vier Hauptbereichen (die nachstehend angegeben sind) berücksichtigt werden.

Quellenauswahl

50. Der erste Bewertungsbereich betrifft die Prüfung der Quellen, die für das Trainieren des KI-Modells ausgewählt wurden. Dazu gehört eine von den Aufsichtsbehörden vorgenommene Bewertung der Vorkehrungen, die gegebenenfalls getroffen wurden, um die Erhebung personenbezogener Daten zu vermeiden oder einzuschränken; dazu gehört unter anderem die Bewertung (i) der Geeignetheit der Auswahlkriterien; (ii) der Relevanz und Angemessenheit der ausgewählten Quellen im Hinblick auf den/die vorgesehenen Zweck(e); sowie (iii) ob ungeeignete Quellen ausgeschlossen wurden.

Datenaufbereitung und Datenminimierung

51. Der zweite Bewertungsbereich betrifft die Aufbereitung der Daten für die Trainingsphase. Die Aufsichtsbehörden sollten insbesondere Folgendes untersuchen: (i) ob in Erwägung gezogen wurde, anonyme und/oder pseudonymisierte personenbezogene Daten zu verwenden; und (ii) falls beschlossen wurde, keine solchen Vorkehrungen zu treffen: die Gründe für diese Entscheidung, wobei der vorgesehene Zweck zu berücksichtigen ist; (iii) die Datenminimierungsstrategien und -techniken, die eingesetzt wurden, um die Menge der im Trainingsverfahren verwendeten personenbezogenen Daten zu beschränken; sowie (iv) (gegebenenfalls) die Datenfilterungsprozesse, die, bevor das Modell trainiert wurde, eingesetzt wurden, um irrelevante personenbezogene Daten zu entfernen.

Methodenauswahl für das Training

52. Der dritte Bewertungsbereich betrifft die Auswahl robuster Methoden für die Entwicklung des KI-Modells. Die Aufsichtsbehörden sollten die ausgewählten Methoden im Hinblick darauf bewerten, ob sie geeignet sind, die Identifizierbarkeit erheblich zu reduzieren oder zu beseitigen; dazu gehört insbesondere die Bewertung: (i) ob die betreffende Methodik Regularisierungsmethoden einsetzt, um die Generalisierung des Modells zu verbessern und Überanpassung zu reduzieren; sowie, und dies ist besonders wichtig, (ii) ob der Verantwortliche geeignete und wirksame Techniken zum Schutz der Privatsphäre eingesetzt hat (z. B. Differential Privacy).

Vorkehrungen in Bezug auf die Modellausgaben

53. Der letzte Bewertungsbereich betrifft dem KI-Modell selbst hinzugefügte Methoden oder Vorkehrungen, die für das Risiko, dass jemand personenbezogene Daten aus dem Modell extrahiert, indem er direkt darauf zugreift, möglicherweise nicht von Belang sind, die jedoch die Wahrscheinlichkeit reduzieren, dass jemand durch Anfragen personenbezogene Daten aus Trainingsdaten erlangt.

3.2.2.2 Prüfung des KI-Modells

54. Aufsichtsbehörden sollten, wenn sie prüfen, ob das KI-Modell in Bezug auf die Anonymisierung robust gestaltet ist, zunächst sicherstellen, dass die Konzeption planmäßig entwickelt wurde und wirksamer Engineering Governance unterliegt. Die Aufsichtsbehörden sollten bewerten, ob die Verantwortlichen dokumentenbasierte (interne oder externe) Audits durchgeführt haben, im Zuge derer die ausgewählten Vorkehrungen und die mit diesen erzielbare Reduzierung der Identifizierungswahrscheinlichkeit bewertet wurden. Dies könnte auch eine Analyse der Code-Review-Berichte sowie eine theoretische Analyse beinhalten, in der die Geeignetheit der Vorkehrungen dokumentiert ist, die für das betreffende Modell ausgewählt wurden, um die Wahrscheinlichkeit der Re-Identifizierung zu reduzieren.

3.2.2.3 KI-Modell-Tests und Angriffsresistenz

55. Im letzten Schritt sollten die Aufsichtsbehörden Umfang, Häufigkeit, Anzahl und Qualität der Tests berücksichtigen, die der Verantwortliche am Modell durchgeführt hat. Insbesondere sollten die Aufsichtsbehörden berücksichtigen, dass erfolgreich verlaufene Tests, die ein breites Spektrum bekannter, dem Stand der Technik entsprechender Angriffe abdecken, lediglich über die Resistenz gegenüber diesen Angriffen Aufschluss geben. Zum Datum dieser Stellungnahme zählen dazu unter anderem strukturierte Tests zur Prüfung der Resistenz gegen: (i) Attribute und Membership Inference; (ii) Exfiltration; (iii) Regurgitation von Trainingsdaten; (iv) Model Inversion; bzw. (v) Reconstruction Attacks.

3.2.2.4 Dokumentation

56. Gemäß den Artikeln 5, 24, 25 und 30 DSGVO sowie – in Fällen eines voraussichtlich hohen Risikos für die Rechte und Freiheiten betroffener Personen – Artikel 35 DSGVO sind die Verantwortlichen gehalten, ihre Verarbeitungsvorgänge angemessen zu dokumentieren. Dies gilt auch für jede Verarbeitung, im Zuge derer ein KI-Modell trainiert wird, selbst wenn das Ziel der Verarbeitung die Anonymisierung ist. Die Aufsichtsbehörden sollten die Dokumentation und etwaige regelmäßige Bewertungen der Folgerisiken für die von den Verantwortlichen durchgeführten Verarbeitungen berücksichtigen, da diese Grundvoraussetzungen für den Nachweis sind, dass keine Verarbeitung personenbezogener Daten stattfindet.

57. **Nach Auffassung des EDSA sollten die Aufsichtsbehörden, die Dokumentation berücksichtigen, wann immer es darum geht, die behauptete Anonymität eines bestimmten KI-Modells zu bewerten. Der EDSA merkt an, dass die Aufsichtsbehörde, falls sie, nachdem sie die behauptete Anonymität unter anderem im Licht der Dokumentation bewertet hat, nicht bestätigen kann, dass wirksame Maßnahmen zur Anonymisierung des KI-Modells ergriffen wurden, in Betracht ziehen könnte, festzustellen, dass die Rechenschaftspflichten des Verantwortlichen aus Artikel 5 Absatz 2 DSGVO nicht erfüllt wurden. Deshalb sollte auch die Einhaltung anderer Bestimmungen der DSGVO geprüft werden.**

58. Idealerweise sollten die Aufsichtsbehörden überprüfen, ob die Dokumentation des Verantwortlichen Folgendes beinhaltet:

    a. Informationen über die Datenschutz-Folgenabschätzungen (DSFA), einschließlich Bewertungen und Entscheidungen, durch die festgestellt wurde, dass keine DSFA erforderlich war;

    b. Ratschläge oder Feedback des Datenschutzbeauftragten („**DSB**“) (soweit ein DSB bestellt wurde oder hätte bestellt werden müssen);

c. Informationen über die technischen und organisatorischen Vorkehrungen, die bei der Gestaltung des KI-Modells getroffen wurden, um die Identifizierungswahrscheinlichkeit zu reduzieren, einschließlich des Bedrohungsmodells und der Risikobewertungen, auf denen diese Vorkehrungen beruhen. Dies sollte spezifische Maßnahmen für jede Quelle der Trainingsdatensätze beinhalten, einschließlich der URLs relevanter Quellen und Beschreibungen der ergriffenen (oder bereits von Datensatz-Drittanbietern ergriffenen) Maßnahmen;

d. die in allen Phasen des Modell-Lebenszyklus ergriffenen technischen und organisatorischen Maßnahmen, die dazu beitrugen oder mittels derer überprüft wurde, dass das Modell keine personenbezogenen Daten enthält;

e. die Dokumentation für den Nachweis der theoretischen Resistenz des KI-Modells gegen Techniken zur Re-Identifizierung wie auch Kontrollen, die darauf ausgelegt sind, Erfolg und Auswirkungen der wichtigsten Angriffsarten (Regurgitation, Membership Inference Attacks, Exfiltration usw.) zu begrenzen oder zu bewerten. Dies kann insbesondere Folgendes beinhalten: (i) das Verhältnis zwischen der Menge der Trainingsdaten und der Anzahl der Parameter im Modell, einschließlich der Analyse der Auswirkungen auf das Modell [38]; (ii) Messgrößen für die Wahrscheinlichkeit der Re-Identifizierung, bezogen auf den Stand der Technik; (iii) Berichte darüber, wie das Modell getestet wurde (von wem, wann, wie und in welchem Umfang), sowie (iv) die Testergebnisse;

f. die dem/den das Modell einsetzenden Verantwortlichen und/oder den betroffenen Personen zur Verfügung gestellte Dokumentation, insbesondere die Dokumentation über die zur Reduzierung der Identifizierungswahrscheinlichkeit ergriffenen Maßnahmen sowie über die möglichen Restrisiken.

3.3 Geeignetheit des berechtigten Interesses als Rechtsgrundlage für die Verarbeitung personenbezogener Daten im Zusammenhang mit Entwicklung und Einsatz von KI-Modellen

59. Zur Beantwortung der Fragen 2 und 3 des Ersuchens wird der EDSA zunächst allgemeine Ausführungen machen zu einigen wichtigen Aspekten, die von den Aufsichtsbehörden – unabhängig von der Rechtsgrundlage für die Verarbeitung – berücksichtigt werden sollten, wenn sie bewerten, auf welche Weise die Verantwortlichen die Einhaltung der DSGVO im Zusammenhang mit KI-Modellen nachweisen können. Sodann wird der EDSA, aufbauend auf den Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO[39], die zur Bewertung des berechtigten Interesses erforderliche dreistufige Prüfung im Zusammenhang mit der Entwicklung und dem Einsatz von KI-Modellen betrachten.

[38] Ricciato F., *A Cautionary Reflection on (Pseudo-)Synthetic Data from Deep Learning on Personal Data*, Privacy in Statistical Databases conference (PSD 2024), Antibes, Frankreich, September 2024, Folien abrufbar unter: https://cros.ec.europa.eu/system/files/2024-10/20240926_PSD2024_Ricciato_v6_1.pdf und Belkin M., Hsu D., Ma S., u. Mandal S. (2019), *Reconciling modern machine-learning practice and the classical bias-variance trade-off*. Proceedings of the National Academy of Sciences, 24. Juli 2019, 116(32) 15849-15854, abrufbar unter: https://www.pnas.org/doi/10.1073/pnas.1903070116

[39] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024.

3.3.1 Allgemeine Bemerkungen

60. Der EDSA erinnert daran, dass die DSGVO keine Hierarchie der verschiedenen in Artikel 6 Absatz 1 DSGVO aufgeführten Rechtsgrundlagen vorsieht[40].

61. In Artikel 5 DSGVO sind die Datenschutzgrundsätze für die Verarbeitung personenbezogener Daten niedergelegt. Der EDSA zeigt auf, welche der Datenschutzgrundsätze für diese Stellungnahme von Bedeutung sind und deshalb von den Aufsichtsbehörden mindestens bei der Bewertung spezifischer KI-Modelle berücksichtigt werden sollten, aber auch die relevantesten Anforderungen, die sich im Hinblick auf den Gegenstand dieser Stellungnahme aus anderen Bestimmungen der DSGVO ergeben.

62. **Grundsatz der Rechenschaftspflicht** (Artikel 5 Absatz 2 DSGVO) – Nach diesem Grundsatz ist der Verantwortliche für die Einhaltung der DSGVO verantwortlich und er muss deren Einhaltung nachweisen können. Diesbezüglich sollten die Rollen und Verantwortlichkeiten derjenigen, die im Zusammenhang mit der Entwicklung oder dem Einsatz eines KI-Modells personenbezogene Daten verarbeiten, einer Bewertung unterzogen werden, bevor die Verarbeitung erfolgt, damit die Pflichten der Verantwortlichen oder gemeinsam Verantwortlichen sowie (gegebenenfalls) der Auftragsverarbeiter gleich zu Beginn feststehen.

63. **Rechtmäßigkeit, Verarbeitung nach Treu und Glauben, Transparenz** (Artikel 5 Absatz 1 Buchstabe a DSGVO) – Für die Bewertung der Rechtmäßigkeit der Verarbeitung im Zusammenhang mit KI-Modellen hält es der EDSA im Hinblick auf Artikel 6 Absatz 1 DSGVO für sinnvoll, zwischen den verschiedenen Phasen der Verarbeitung personenbezogener Daten zu unterscheiden[41]. Der Grundsatz von Treu und Glauben (Fairness), der mit dem Transparenzgrundsatz in engem Zusammenhang steht, erfordert, dass personenbezogene Daten nicht mittels unfairer Methoden oder in irreführender Weise oder auf eine Weise, die „für die betroffene Person in nicht gerechtfertigter Weise schädlich, widerrechtlich diskriminierend, unerwartet oder irreführend“[42] ist, verarbeitet werden. Angesichts der Komplexität der Technologien, um die es geht, sollten die Informationen über die Verarbeitung personenbezogener Daten in KI-Modellen in zugänglicher, verständlicher und nutzerfreundlicher Weise bereitgestellt werden [43] . Zur Transparenz hinsichtlich der Verarbeitung personenbezogener Daten gehört insbesondere die Einhaltung der in den Artikeln 12 bis 14 DSGVO geregelten Informationspflichten[44], wobei im Fall der automatisierten Entscheidungsfindung einschließlich Profiling auch aussagekräftige Informationen über die involvierte Logik sowie die Tragweite und die angestrebten Auswirkungen einer derartigen Verarbeitung für die betroffene Person [45] erforderlich sind. Da in der Entwicklungsphase von KI-Modellen unter Umständen große Datenmengen aus allgemein

[40] Ebd., Randnummer 1.
[41] EDPB Report of the work undertaken by the ChatGPT Taskforce (Bericht des EDSA über die von der ChatGPT Taskforce unternommene Arbeit), angenommen am 23. Mai 2024, Randnummer 14.
[42] EDPB Report of the work undertaken by the ChatGPT Taskforce (Bericht des EDSA über die von der ChatGPT Taskforce unternommene Arbeit), angenommen am 23. Mai 2024, Randnummer 23; EDSA, Leitlinien 4/2019 zu Artikel 25 Datenschutz durch Technikgestaltung und durch datenschutzfreundliche Voreinstellungen, Version 2.0, angenommen am 20. Oktober 2020, Randnummer 69; Artikel-29-Datenschutzgruppe, Leitlinien für Transparenz gemäß der Verordnung 2016/679, zuletzt überarbeitet und angenommen am 11. April 2018, vom EDSA gebilligt am 25. Mai 2018, Randnummer 2.
[43] Artikel-29-Datenschutzgruppe, Leitlinien für Transparenz gemäß der Verordnung 2016/679, zuletzt überarbeitet und angenommen am 11. April 2018, vom EDSA gebilligt am 25. Mai 2018, Randnummer 5.
[44] Vgl. auch Erwägungsgrund 39 der DSGVO, wo es heißt, dass „[f]ür natürliche Personen … Transparenz dahin gehend bestehen [sollte], dass sie betreffende personenbezogene Daten erhoben, verwendet, eingesehen oder anderweitig verarbeitet werden und in welchem Umfang die personenbezogenen Daten verarbeitet werden und künftig noch verarbeitet werden …“.
[45] Artikel 13 Absatz 2 Buchstabe f DSGVO und Artikel 14 Absatz 2 Buchstabe g DSGVO.

zugänglichen Quellen erhoben werden (beispielsweise durch Web-Scraping-Techniken), ist die Inanspruchnahme des in Artikel 14 Absatz 5 Buchstabe b DSGVO vorgesehenen Ausnahmetatbestands strikt auf Fälle beschränkt, in denen die Anforderungen dieser Bestimmung in vollem Umfang erfüllt sind[46].

64. **Grundsätze der Zweckbindung und der Datenminimierung** (Artikel 5 Absatz 1 Buchstaben b und c der DSGVO) – Gemäß dem Grundsatz der Datenminimierung erfordern die Entwicklung und der Einsatz von KI-Modellen, dass die personenbezogenen Daten dem Zweck angemessen, erheblich und notwendig sein sollten. Dabei können personenbezogene Daten verarbeitet werden, um Risiken in Bezug auf potenzielle Verzerrungen und Fehler zu vermeiden, sofern dies im Rahmen des Zweckfestlegung klar und konkret angegeben wird und die personenbezogenen Daten für den Zweck erforderlich sind (z. B. weil dies nicht auf effektive Weise durch die Verarbeitung anderer Daten, beispielsweise synthetischer oder anonymisierter Daten, möglich ist)[47]. Die WP 29 hat bereits betont, dass „[d]er Zweck der Erhebung … klar und genau angegeben sein muss …“[48]. Für die Bewertung, ob der verfolgte Zweck legitim, festgelegt und eindeutig ist und ob die Verarbeitung mit dem Grundsatz der Datenminimierung in Einklang steht, sollte zunächst die in Rede stehende Verarbeitungstätigkeit ermittelt werden. Dabei ist zu beachten, dass es sich bei den verschiedenen Stadien der Entwicklungs- oder Einsatzphasen um dieselben oder um verschiedene Verarbeitungstätigkeiten handeln kann, unter Umstände mit einer Reihe aufeinanderfolgender Verantwortlicher oder gemeinsam Verantwortlicher. In einigen Fällen ist es möglich, den Zweck, der mit dem Einsatz eines KI-Modells verfolgt werden wird, bereits in einem frühen Entwicklungsstadium zu bestimmen. Doch selbst wenn dies nicht der Fall sein sollte, sollte der Zusammenhang, in dem der Einsatz erfolgt, doch schon in gewissem Umfang bekannt sein; deshalb sollte in Betracht gezogen werden, was sich aus diesem Zusammenhang über den Zweck der Entwicklung ableiten lässt. Wenn sie den Verarbeitungszweck für ein bestimmtes Entwicklungsstadium prüfen, sollten die Aufsichtsbehörden erwarten, dass die Verantwortlichen einen gewissen Grad an Detailinformationen liefern und erklären, was sich aus diesen Details in Bezug auf den Verarbeitungszweck ergibt. In Betracht kommen beispielsweise Informationen über die Art des entwickelten KI-Modells, dessen voraussichtliche Funktionsweise sowie jeglicher sonstige relevante Kontext, der in dem betreffenden Stadium bereits bekannt ist. Was den Einsatzkontext angeht, kämen beispielsweise Informationen darüber in Betracht, ob das Modell für den internen Einsatz entwickelt wird, ob der Verantwortliche das Modell nach der Entwicklung an Dritte zu verkaufen oder zu vertreiben beabsichtigt, wie auch, ob das Modell in erster Linie für Forschungs- oder für gewerbliche Zwecke eingesetzt werden soll.

65. **Recht der betroffenen Person** (Kapitel III der DSGVO) – Dessen ungeachtet müssen die Aufsichtsbehörden sicherstellen, dass die Verantwortlichen bei der Entwicklung und dem Einsatz von KI-Modellen alle Rechte der betroffenen Personen wahren; insoweit erinnert der EDSA daran, dass, wenn sich Verantwortliche auf die Rechtsgrundlage des berechtigten Interesses stützen, stets auch ein

[46] EDPB Report of the work undertaken by the ChatGPT Taskforce (Bericht des EDSA über die von der ChatGPT Taskforce unternommene Arbeit), angenommen am 23. Mai 2024, Randnummer 27.
[47] Darüber hinaus sieht Artikel 10 Absatz 5 der KI-Verordnung zur Erkennung und Korrektur von Verzerrungen im Zusammenhang mit Hochrisiko-KI-Systemen spezifische Regeln für die Verarbeitung besonderer Kategorien personenbezogener Daten vor.
[48] Artikel-29-Datenschutzgruppe, Stellungnahme 03/2013 zur Zweckbindung (WP203), S. 15-16.

Widerspruchsrecht gemäß Artikel 21 DSGVO besteht und dass dieses Widerspruchsrecht sicherzustellen ist[49].

3.3.2 Erwägungen in Bezug auf die dreistufige Prüfung der berechtigten Interessen im Zusammenhang mit der Entwicklung und dem Einsatz von KI-Modellen

66. Für die Bewertung, ob eine bestimmte Verarbeitung personenbezogener Daten auf Artikel 6 Absatz 1 Buchstabe f DSGVO gestützt werden kann, sollten die Aufsichtsbehörden überprüfen, ob die Verantwortlichen sorgfältig geprüft und dokumentiert haben, ob die drei folgenden kumulativen Voraussetzungen erfüllt sind: (i) dass von dem für die Verarbeitung Verantwortlichen oder von einem Dritten ein berechtigtes Interesse wahrgenommen wird; (ii) dass die Verarbeitung der personenbezogenen Daten zur Verwirklichung des berechtigten Interesses erforderlich ist; und (iii) dass die Interessen oder Grundrechte und Grundfreiheiten der betroffenen Personen gegenüber dem berechtigten Interesse des Verantwortlichen oder eines Dritten nicht überwiegen[50].

3.3.2.1 Erster Prüfungsschritt – Wahrung der berechtigten Interessen des Verantwortlichen oder eines Dritten

67. Ein Interesse ist das umfassendere Interesse oder der Nutzen, den ein für die Verarbeitung Verantwortlicher oder ein Dritter an einer bestimmten Verarbeitungstätigkeit haben kann[51]. Während sowohl in der DSGVO als auch vom EuGH verschiedene Interessen als legitim anerkannt werden[52], sollte die Bewertung, ob es sich um ein berechtigtes Interesse handelt, doch das Ergebnis einer Einzelfallprüfung sein.

68. In seinen Leitlinien über berechtigte Interessen[53] erinnert der EDSA daran, dass ein Interesse als legitim angesehen werden kann, wenn die folgenden drei kumulativen Kriterien erfüllt sind:

a. Das Interesse ist rechtmäßig[54];

---

[49] Artikel 21 DSGVO bestimmt für den Fall, dass eine betroffene Person aus Gründen, die sich aus ihrer besonderen Situation ergeben, Widerspruch gegen die Verarbeitung sie betreffender personenbezogener Daten einlegt, dass der Verantwortliche die personenbezogenen Daten nicht mehr verarbeiten darf, es sei denn, er kann zwingende schutzwürdige Gründe für die Verarbeitung nachweisen, die die Interessen, Rechte und Freiheiten der betroffenen Person überwiegen, oder die Verarbeitung dient der Geltendmachung, Ausübung oder Verteidigung von Rechtsansprüchen. Die Aufsichtsbehörden müssen dazu zwei Aspekte prüfen: ob der Verantwortliche solche schwerer wiegenden zwingenden schutzwürdigen Gründe nachweisen kann und ob das Widerspruchsrecht ausgeübt werden kann.

[50] EuGH, Urteil vom 4. Juli 2023, Rechtssache C-252/21, *Meta/Bundeskartellamt* (ECLI:EU:C:2023:537), Randnummer 106; EuGH, Urteil vom 11. Dezember 2019, Rechtssache C-708/18, *Asociația de Proprietari bloc M5A-ScaraA* (ECLI:EU:C:2019:1064), Randnummer 40. Vgl. auch EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummern 12 ff. In diesen Leitlinien wird daran erinnert, dass„*[d]ie Bewertung … zu Beginn der Verarbeitung unter Einbeziehung des Datenschutzbeauftragten (DSB) (falls benannt) erfolgen und von dem für die Verarbeitung Verantwortlichen im Einklang mit dem in Artikel 5 Absatz 2 DSGVO festgelegten Grundsatz der Rechenschaftspflicht dokumentiert werden [sollte]*“.

[51] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummer 14.

[52] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummer 16.

[53] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummer 17.

[54] EuGH, Urteil vom 4. Oktober 2024, Rechtssache C-621/22, *Koninklijke Nederlandse Lawn Tennisbond* (ECLI:EU:C:2024:857), Randnummer 49, wo der EuGH hervorgehoben hat, dass ein berechtigtes Interesse nicht

b. das Interesse ist klar und präzise formuliert und

c. das Interesse ist real und gegenwärtig und darf nicht spekulativ sein.

69. Vorbehaltlich der beiden anderen Schritte, die zur Prüfung berechtigter Interessen erforderlich sind, könnten im Zusammenhang mit KI-Modellen die folgenden Beispiele berechtigte Interessen darstellen: (i) die Entwicklung eines Konversationsagenten, der Nutzern Hilfe leistet; (ii) die Entwicklung eines KI-Systems, das betrügerische Inhalte oder Verhaltensweisen aufdeckt, und (iii) die Verbesserung der Bedrohungserkennung in einem Informationssystem.

3.3.2.2 Zweiter Prüfungsschritt – Prüfung der Erforderlichkeit der Verarbeitung zur Verwirklichung des berechtigten Interesses

70. Im zweiten Prüfungsschritt ist zu bestimmen, ob die Verarbeitung personenbezogenen Daten zur Verwirklichung des berechtigten Interesses erforderlich ist[55] („Erforderlichkeitsprüfung“).

71. In Erwägungsgrund 39 DSGVO wird klargestellt, dass „[p]ersonenbezogene Daten … nur [sollten] verarbeitet werden dürfen, wenn der Zweck der Verarbeitung nicht in zumutbarer Weise durch andere Mittel erreicht werden kann“. Im Einklang mit der Rechtsprechung des EuGH und den früheren Leitlinien des EDSA ist die Voraussetzung der Erforderlichkeit der Verarbeitung im Licht der Grundrechte und Grundfreiheiten der betroffenen Personen in Verbindung mit dem in Artikel 5 Absatz 1 Buchstabe c DSGVO verankerten Grundsatz der Datenminimierung zu prüfen[56].

72. Die vom EuGH vorgesehene Methode berücksichtigt sowohl den Kontext der Verarbeitung als auch die Auswirkungen auf den Verantwortlichen und die betroffenen Personen. Die Erforderlichkeitsprüfung

gesetzwidrig sein kann. Insoweit weist der EDSA darauf hin, dass bei der Beurteilung der Rechtmäßigkeit eines Interesses die gesetzlichen Rahmenwerke berücksichtigt werden sollten. Vgl. zum Beispiel: Artikel 26 Absatz 3 und Artikel 28 der Verordnung (EU) 2022/2065 des Europäischen Parlaments und des Rates vom 19. Oktober 2022 über einen Binnenmarkt für digitale Dienste und zur Änderung der Richtlinie 2000/31/EG (Gesetz über digitale Dienste) („GdD“) im Hinblick auf das Verbot an Minderjährige gerichteter zielgruppenspezifischer Werbung; Artikel 5 Absätze 1 und 2 der KI-Verordnung über verbotene Praktiken im KI-Bereich (manipulative Praktiken und unterschwellige Beeinflussung außerhalb des Bewusstseins einer Person); Verarbeitung unter Verletzung von Rechten des geistigen Eigentums und der Bestimmungen der Richtlinie (EU) 2019/790 über das Urheberrecht und die verwandten Schutzrechte im digitalen Binnenmarkt.

[55] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummern 28-30.

[56] EuGH, Urteil vom 4. Juli 2023, Rechtssache C-252/21, *Meta/Bundeskartellamt* (ECLI:EU:C:2023:537), Randnummern 108 und 109, auch unter Bezugnahme auf EuGH, Urteil vom 11. Dezember 2019, Rechtssache C-708/18, *Asociaţia de Proprietari bloc M5A-ScaraA* (ECLI:EU:C:2019:1064), Randnummer 48; EuGH, Urteil vom 9. November 2010, verbundene Rechtssachen C-92/09 und C-93/09, *Volker und Markus Schecke* (ECLI:EU:C:2010:662), Randnummern 85 und 86; EuGH, Urteil vom 22. Juni 2021, Rechtssache C-439/19, *Latvijas Republikas Saeima* (ECLI:EU:C:2021:504), Randnummern 98, 109, 110, 113. Siehe beispielsweise: EDSA, Leitlinien 3/2019 zur Verarbeitung personenbezogener Daten durch Videogeräte, Version 2.0, angenommen am 29. Januar 2020, Randnummern 24-26 und 73; EDSA, Leitlinien 2/2019 für die Verarbeitung personenbezogener Daten gemäß Artikel 6 Absatz 1 Buchstabe b DSGVO im Zusammenhang mit der Erbringung von Online-Diensten für betroffene Personen, Version 2.0, angenommen am 8. Oktober 2019, Randnummern 23-25; EDSA, Stellungnahme Nr. 11/2024 zum Einsatz von Gesichtserkennung zur Straffung der Fluggastströme, Version 1.1, angenommen am 23. Mai 2024, Randnummer 27.

umfasst also zwei Elemente: (i) ob die Verarbeitungstätigkeit die Zweckverwirklichung ermöglicht[57]; und (ii) ob es keinen weniger eingreifenden Weg gibt, diesen Zweck zu verwirklichen[58].

73. Beispielsweise ist es, je nach den Umständen des Einzelfalls, erforderlich, die vorgesehene Menge der in das KI-Modell eingegangenen personenbezogenen Daten im Licht weniger stark in die Privatsphäre eingreifender Alternativen zu bewerten, von denen angemessenerweise Gebrauch gemacht werden könnte und die den Zweck der verfolgten berechtigten Interessen genauso wirksam erzielen. Soweit es möglich ist, den Zweck durch ein KI-Modell zu verfolgen, das ohne die Verarbeitung personenbezogener Daten auskommt, sollte die Verarbeitung personenbezogener Daten deshalb für nicht erforderlich befunden werden. Von besonderer Relevanz ist dies für die Entwicklung von KI-Modellen. Für die Bewertung, ob die Erforderlichkeitsvoraussetzung erfüllt ist, sollten die Aufsichtsbehörden insbesondere auf die Menge der verarbeiteten personenbezogenen Daten achten, sowie darauf, ob diese in angemessenem Verhältnis zu den im Einzelfall verfolgten berechtigten Interessen steht, was auch im Licht des Grundsatzes der Datenminimierung zu beurteilen ist.

74. Die Erforderlichkeitsprüfung sollte auch den größeren Zusammenhang der beabsichtigten Verarbeitung personenbezogener Daten berücksichtigen. Ob es Mittel gibt, die weniger stark in die Grundfreiheiten und Grundrechte der betroffenen Personen eingreifen, kann davon abhängen, ob der Verantwortliche in direkter Beziehung zu den betroffenen Personen steht (im Fall von ihm selbst erhobener Daten) oder nicht (im Fall von Dritten erhobener Daten). Der EuGH hat einige der Erwägungen aufgezeigt, die bei der Prüfung der Erforderlichkeit der Verarbeitung selbst erhobener Daten für die Zwecke eines oder mehrerer verfolgter berechtigter Interessen anzustellen sind (wobei es allerdings um die Offenlegung derartiger Daten gegenüber Dritten ging)[59].

75. Auch die Umsetzung technischer Garantien zum Schutz personenbezogener Daten kann dazu beitragen, die Anforderungen der Erforderlichkeitsprüfung zu erfüllen. Darunter fiele beispielsweise, Maßnahmen wie diejenigen in Abschnitt 3.2.2 auf solche Weise umzusetzen, dass zwar keine Anonymisierung erzielt wird, dass jedoch die Identifizierung betroffener Personen weniger leicht

---

[57] Vgl. EuGH, Urteil vom 16. Dezember 2008, Rechtssache C-524/06, *Huber / Bundesrepublik Deutschland* (ECLI:EU:C:2008:724), Randnummer 66. Vgl. auch in derselben Rechtssache Nummer 16 der Schlussanträge des Generalanwalts Poiares Maduro (Rechtssache *Huber / Bundesrepublik Deutschland*, C-524/06, (ECLI:EU:C:2008:194)), Randnummer 16, wo er ausführt: „*dass hier richtigerweise auf die Effektivität abzustellen ist und dass die entsprechende Prüfung durch das nationale Gericht vorzunehmen ist, das sich die Frage stellen muss, ob es andere Möglichkeiten der Datenverarbeitung gibt, mit denen die Zuwanderungsbehörden die aufenthaltsrechtlichen Bestimmungen vollziehen könnten. Bejaht es diese Frage, sollte die zentrale Speicherung und Verarbeitung der Daten ausländischer Unionsbürger für rechtswidrig erklärt werden. Das alternative System muss nicht unbedingt das absolut effektivste oder geeignetste sein, es reicht aus, dass es angemessen funktioniert. Mit anderen Worten ist selbst dann, wenn das Zentralregister effektiver oder praktischer oder anwenderfreundlicher ist als andere Möglichkeiten (wie etwa die dezentralisierten Register bei den Gemeinden), Letzteren klar der Vorzug zu geben, wenn sie dazu genutzt werden können, den aufenthaltsrechtlichen Status ausländischer Unionsbürger anzugeben*“.

[58] Vgl. EuGH, Urteil vom 27. September 2017, Rechtssache C-73/16, *Peter Puškár* (ECLI:EU:C:2017:725), Randnummer 113: *„Das vorlegende Gericht hat somit zu prüfen, ob die Erstellung der streitigen Liste und die Aufnahme des Namens der betroffenen Personen in diese geeignet sind, die damit verfolgten Ziele zu verwirklichen, und ob es nicht andere, mildere Mittel zur Erreichung dieser Ziele gibt.“;* Vgl. auch z. B. Schlussanträge des Generalanwalts Rantos in der Rechtssache C-252/21, *Meta/Bundeskartellamt* (ECLI:EU:C:2022:704), Randnummer 61, wo es heißt: *„… Daher muss eine enge Verbindung zwischen der Verarbeitung und dem wahrgenommenen Interesse bestehen, wenn es keine den Schutz personenbezogener Daten weniger beeinträchtigenden Alternativen gibt, weil es nicht ausreicht, dass die Verarbeitung für den für die Verarbeitung Verantwortlichen lediglich von Nutzen ist.“*

[59] EuGH, Urteil vom 4. Oktober 2024, Rechtssache C-621/22, *Koninklijke Nederlandse Lawn Tennisbond* (ECLI:EU:C:2024:857), Randnummern 51-53.

möglich ist. Der EDSA merkt an, dass einige dieser Maßnahmen, auch wenn sie nicht zur Einhaltung der DSGVO erforderlich sind, zusätzliche Garantien darstellen können; vgl. dazu die näheren Ausführungen im Unterabschnitt „Risikomindernde Maßnahmen“ des Abschnitts 3.3.2.3[60].

3.3.2.3 Dritter Prüfungsschritt – Abwägungsprüfung

76. Dieser dritte Schritt in der Prüfung der berechtigten Interessen ist die „**Abwägung**“ (in diesem Dokument auch als „**Abwägungsprüfung**“ bezeichnet) [61] . Dieser Schritt besteht darin, die verschiedenen einander gegenüberstehenden Rechte und Interessen, um die es in der Sache geht, zu erkennen und zu beschreiben [62] , d. h. auf der einen Seite die Interessen, Grundfreiheiten und Grundrechte der betroffenen Personen und auf der anderen Seite die Interessen des Verantwortlichen oder Dritten. Für den Nachweis, dass die berechtigten Interessen als Rechtsgrundlage für die in Rede stehenden Verarbeitungstätigkeiten dienen können, sollten sodann die besonderen Umstände des Einzelfalls berücksichtigt werden[63].

Interessen, Grundrechte und Grundfreiheiten der betroffenen Personen

77. Artikel 6 Absatz 1 Buchstabe f DSGVO bestimmt, dass der Verantwortliche, wenn er die verschiedenen Aspekte im Rahmen der Abwägungsprüfung bewertet, die Interessen, Grundrechte und Grundfreiheiten der betroffenen Personen berücksichtigen muss. Die Interessen der betroffenen Personen sind diejenigen, die von der in Rede stehenden Verarbeitung betroffen sein können. Im Zusammenhang mit der Entwicklungsphase eines KI-Modells sind dies unter Umständen, wobei dies keine abschließende Aufzählung ist, das Interesse an Selbstbestimmung und daran, die Kontrolle über die eigenen personenbezogenen Daten (z. B. die für die Modellentwicklung erhobenen Daten) zu behalten. Im Zusammenhang mit dem Einsatz eines KI-Modells handelt es sich bei den Interessen der betroffenen Personen unter Umständen, wobei dies keine abschließende Aufzählung ist, um das Interesse daran, die Kontrolle über die eigenen personenbezogenen Daten zu behalten (z. B. in Bezug auf die Daten, die verarbeitet werden, nachdem das Modell eingesetzt wurde), um finanzielle Interessen (z. B. wenn die betroffene Person das KI-Modell zur Einnahmenerzielung oder im Rahmen ihrer beruflichen Tätigkeit verwendet), um persönlichen Nutzen (z. B. wenn das KI-Modell besseren Zugang zu gewissen Diensten bietet) oder um sozioökonomische Interessen (z. B. wenn das KI-Modell besseren Zugang zum Gesundheitswesen ermöglicht oder die Ausübung von Grundrechten erleichtert, beispielsweise die Ausübung des Grundrechts auf Bildung)[64].

78. Je genauer ein Interesse im Hinblick auf den vorgesehenen Zweck der Verarbeitung definiert ist, desto besser wird es möglich sein, die tatsächlichen Gegebenheiten in Bezug auf die für die Abwägung zu berücksichtigenden Nutzen und Risiken klar zu verstehen.

79. Entwicklung und Einsatz von KI-Modellen können zu schwerwiegenden Risiken im Hinblick auf die durch die Charta der Grundrechte der Europäischen Union („**EU-Charta**“) geschützten Grundrechte und Grundfreiheiten der betroffenen Personen führen; beispielsweise, wobei dies keine abschließende

[60] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummer 57.
[61] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummern 31-60.
[62] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummer 32.
[63] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummer 32, wo auch Bezug genommen wird auf EuGH, Urteil vom 4. Juli 2023, Rechtssache C-252/21, *Meta/Bundeskartellamt* (ECLI:EU:C:2023:537), Randnummer 110.
[64] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummer 38.

Aufzählung ist, zu Risiken für Rechte wie das Recht auf Achtung des Privat- und Familienlebens (Artikel 7 EU-Charta) und das Recht auf den Schutz personenbezogener Daten (Artikel 8 EU-Charta). Diese Risiken können sich in der Entwicklungsphase ergeben, z. B. wenn personenbezogene Daten gegen den Willen der betroffenen Person oder ohne deren Wissen durch Scraping ausgelesen werden. Derartige Risiken können auch in der Einsatzphase entstehen, wenn etwa personenbezogene Daten durch das Modell (oder als Teil desselben) in gegen die Rechte der betroffenen Personen verstoßender Weise verarbeitet werden oder wenn es möglich ist, sei es versehentlich oder durch gezielte Angriffe (z. B. Membership Inference, Extraction oder Model Inversion), abzuleiten, welche personenbezogenen Daten in der für das maschinelle Lernen verwendeten Datenbank enthalten waren. In derartigen Situationen besteht ein Risiko für die Privatsphäre der betroffenen Personen, deren Daten in der Einsatzphase des KI-Systems möglicherweise angezeigt werden (je nach Art der Daten z. B. Reputationsrisiko, Identitätsdiebstahl oder -betrug, Sicherheitsrisiko).

80. Je nach dem Einzelfall können auch andere Grundrechte berührt sein. So kann beispielsweise eine groß angelegte und wahllose Datenerhebung in der Entwicklungsphase von KI-Modellen dazu führen, dass betroffene Personen das Gefühl haben, überwacht zu werden; zumal sich nur schwer verhindern lässt, dass öffentliche Daten durch Scraping ausgelesen werden. Dies kann dazu führen, dass sich Menschen eine Selbstzensur auferlegen, wodurch wiederum das Risiko entsteht, dass ihr Recht auf Freiheit der Meinungsäußerung (Artikel 11 EU-Charta) untergraben wird. Auch in der Einsatzphase bestehen Risiken für die Freiheit der Meinungsäußerung und Informationsfreiheit, wenn KI-Modelle dazu verwendet werden, die Veröffentlichung von Inhalten betroffener Personen zu blockieren. Zudem kann ein KI-Modell, das gefährdeten Personen unangemessene Inhalte empfiehlt, deren psychische Gesundheit gefährden (Artikel 3 Absatz 1 EU-Charta). In anderen Fällen kann der Einsatz von KI-Modellen auch zu negativen Folgen für das Recht einer Person, zu arbeiten, (Artikel 15 EU-Charta) führen, etwa wenn Stellenbewerbungen mittels eines KI-Modells vorsortiert werden. Gleichermaßen könnte ein KI-Modell das Recht auf Nichtdiskriminierung (Artikel 21 EU-Charta) berühren, wenn es Personen wegen gewisser persönlicher Merkmale (etwa wegen der Staatsangehörigkeit oder des Geschlechts) diskriminiert. Der Einsatz von KI-Modellen kann auch die Sicherheit und den Schutz von Menschen gefährden (z. B. beim böswilligen Einsatz von KI-Modellen) und Risiken für deren körperliche und geistige Unversehrtheit begründen[65].

81. Allerdings kann der Einsatz von KI-Modellen, was gewisse Grundrechte angeht, auch positive Auswirkungen haben: So kann ein Modell z. B. das Recht auf geistige Unversehrtheit (Artikel 3 EU-Charta) unterstützen, wenn schädliche Online-Inhalte mittels KI-Modellen erkannt werden; Modelle können auch den Zugang zu bestimmten wichtigen Diensten ermöglichen oder die Wahrnehmung von Grundrechten, etwa den Zugang zu Informationen (Artikel 11 EU-Charta) oder den Zugang zu Bildung (Artikel 14 EU-Charta), erleichtern.

Auswirkungen der Verarbeitung auf betroffene Personen

82. Werden im Zuge der Entwicklungsphase und Einsatzphase von KI-Modellen personenbezogene Daten verarbeitet, so kann dies unterschiedliche – positive oder negative – Auswirkungen auf betroffene Personen haben[66]. Wenn zum Beispiel eine Verarbeitungstätigkeit für die betroffene Person von Nutzen ist, kann dies bei der Abwägung berücksichtigt werden. Es ist also durchaus möglich, dass eine Aufsichtsbehörde im Hinblick auf einen solchen Nutzen zu dem Ergebnis gelangt, dass die Interessen, Grundrechte und Grundfreiheiten der betroffenen Personen die Interessen des Verantwortlichen oder

[65] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummer 46.
[66] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummer 39.

eines Dritten nicht überwiegen; eine solche Schlussfolgerung kann jedoch stets nur auf Grundlage des Ergebnisses einer alle relevanten Faktoren berücksichtigenden Einzelfallprüfung gezogen werden.

83. Die Auswirkungen der Verarbeitung auf die betroffenen Personen können beeinflusst sein durch (i) die Art der von den Modellen verarbeiteten Daten; (ii) den Kontext der Verarbeitung und (iii) alle weiteren Folgen, die die Verarbeitung haben mag[67].

84. Im Hinblick auf die **Art der verarbeiteten Daten** ist daran zu erinnern, dass – abgesehen von den besonderen Kategorien personenbezogener Daten und Daten über strafrechtliche Verurteilungen und Straftaten, die jeweils gemäß Artikel 9 und Artikel 10 DSGVO zusätzlichen Schutz genießen – die Verarbeitung einiger anderer Kategorien personenbezogener Daten zu schwerwiegenden Folgen für die betroffenen Personen führen kann. In diesem Zusammenhang ist zu bedenken, dass es für die betroffenen Personen schwerwiegende Folgen haben kann, wenn für die Entwicklung und den Einsatz von KI-Modellen gewisse Arten personenbezogener Daten verarbeitet werden, aus denen höchstprivate Informationen (z. B. Finanzdaten oder Geodaten) hervorgehen. In der Einsatzphase können die Folgen, die sich für die betroffene Person aus einer solchen Verarbeitung ergeben, zum Beispiel ihre wirtschaftliche Lage (z. B. Diskriminierung in der Arbeitswelt) und/oder ihr Ansehen (z. B. Verleumdung) betreffen.

85. Was den **Kontext der Verarbeitung** angeht, ist zunächst festzustellen, welche Elemente Risiken für die betroffenen Personen begründen könnten (z. B. die Art und Weise, in der das Modell entwickelt wurde, die Art und Weise, in der das Modell eingesetzt werden kann, und/oder ob geeignete Sicherheitsvorkehrungen zum Schutz der personenbezogenen Daten getroffen wurden). Für die Ermittlung solcher potenziellen Risikoursachen kommt es entscheidend auf die Art des Modells und die vorgesehenen betrieblichen Verwendungen an.

86. Es ist auch erforderlich zu prüfen, wie schwerwiegend diese Risiken für die betroffenen Personen sind. Unter anderem kann in Betracht gezogen werden, wie die personenbezogenen Daten verarbeitet werden (ob sie z. B. mit anderen Datensätzen kombiniert werden), in welchem Umfang sie verarbeitet werden und wie viele personenbezogene Daten verarbeitet werden[68] (z. B. die Gesamtdatenmenge, die Datenmenge je betroffener Person, die Zahl der betroffenen Personen) [69], der Status der betroffenen Person (z. B. Kinder oder andere schutzbedürftige Personen) und deren Beziehung zum Verantwortlichen (z. B. ob es sich bei der betroffenen Person um einen Kunden handelt). So kann zum Beispiel der Einsatz von Web Scraping in der Entwicklungsphase – wenn es an ausreichenden Garantien fehlt – wegen der großen Menge der erhobenen Daten, der großen Anzahl betroffener Personen und der wahllosen Erhebung personenbezogener Daten zu erheblichen Folgen für natürliche Personen führen.

87. Bei der Bewertung der Auswirkungen auf die betroffenen Personen sollten auch die **weiteren Folgen**, die die Verarbeitung haben könnte, berücksichtigt werden. Diese sollten von den Aufsichtsbehörden im Einzelfall und unter Berücksichtigung der besonderen Umstände bewertet werden.

88. Derartige Folgen sind beispielsweise (wobei dies keine abschließende Aufzählung ist) das Risiko der Verletzung von Grundrechten betroffener Personen (vgl. dazu die Risikobeschreibungen im

[67] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummer 32.
[68] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummer 43.
[69] EuGH, Urteil vom 4. Juli 2023, Rechtssache C-252/21, *Meta/Bundeskartellamt* (ECLI:EU:C:2023:537), Randnummer 116.

vorhergehenden Unterabschnitt[70]). Die Wahrscheinlichkeit und der Schweregrad dieser Risiken mögen unterschiedlich ausfallen, und es kann sein, dass sie sich aus einer Verarbeitung personenbezogener Daten ergeben, die zu körperlichen Schäden, Sachschäden oder immateriellen Schäden führen könnte, insbesondere wenn die Verarbeitung Diskriminierung bewirkt[71].

89. Kommt es im Zuge des Einsatzes eines KI-Modells zur Verarbeitung personenbezogener Daten (i) betroffener Personen, deren personenbezogene Daten in dem in der Entwicklungsphase verwendeten Datensatz enthalten waren, wie auch (ii) betroffener Personen, deren personenbezogene Daten in der Einsatzphase verarbeitet werden, so sollten die Aufsichtsbehörden insoweit unterscheiden und bei der Überprüfung der vom Verantwortlichen durchgeführten Abwägung die Risiken für die Interessen, Rechte und Freiheiten für jede dieser Kategorien betroffener Personen gesondert betrachten.

90. **Abschließend ist zu sagen, dass die Prüfung der möglichen weiteren Folgen der Verarbeitung auch deren Eintrittswahrscheinlichkeit berücksichtigen sollte.** Bei der Beurteilung dieser Wahrscheinlichkeit sollten die technischen und organisatorischen Vorkehrungen wie auch die besonderen Umstände des Einzelfalls berücksichtigt werden. Die Aufsichtsbehörden könnten zum Beispiel prüfen, ob Vorkehrungen zur Verhinderung einer etwaigen missbräuchlichen Verwendung des KI-Modells getroffen wurden. Im Falle von KI-Modellen, die für verschiedenste Zwecke einsetzbar sind (beispielsweise bei generativer KI), kann es sich um Vorkehrungen handeln, die die Möglichkeit, die KI-Modelle missbräuchlich zu verwenden, so weit wie möglich einschränken, insbesondere im Hinblick auf die missbräuchliche Verwendung: zur Erstellung von Deepfakes; für Desinformation, Phishing und andere mittels Chatbots begangene Betrugsformen; für manipulative KI / KI-Agenten (insbesondere solche, die anthropomorphisch sind oder irreführende Informationen liefern).

Vernünftige Erwartungen der betroffenen Personen

91. In Erwägungsgrund 47 der DSGVO heißt es, dass*„[a]uf jeden Fall … das Bestehen eines berechtigten Interesses besonders sorgfältig abzuwägen [wäre], wobei auch zu prüfen ist, ob eine betroffene Person zum Zeitpunkt der Erhebung der personenbezogenen Daten und angesichts der Umstände, unter denen sie erfolgt, vernünftigerweise absehen kann, dass möglicherweise eine Verarbeitung für diesen Zweck erfolgen wird. Insbesondere dann, wenn personenbezogene Daten in Situationen verarbeitet werden, in denen eine betroffene Person vernünftigerweise nicht mit einer weiteren Verarbeitung rechnen muss, könnten die Interessen und Grundrechte der betroffenen Person das Interesse des Verantwortlichen überwiegen“*[72].

92. Bei der Abwägungsprüfung spielen die vernünftigen Erwartungen eine wichtige Rolle, nicht zuletzt, weil die für KI-Modelle verwendete Technologie komplex ist und es für betroffene Personen schwer abzusehen sein mag, in wie vielfältiger Weise ein KI-Modell eingesetzt werden kann und welche Datenverarbeitung damit einhergeht[73]. In diesem Zusammenhang können die den betroffenen

[70] Vgl. den obigen Unterabschnitt „Interessen, Grundrechte und Grundfreiheiten der betroffenen Personen“.
[71] Vgl. Abschnitt 2.3 der EDSA-Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024. Vgl. auch die weiteren Beispiele in Erwägungsgrund 75 der DSGVO.
[72] Vgl. auch EuGH, Urteil vom 4. Juli 2023, Rechtssache C-252/21, *Meta/Bundeskartellamt* (ECLI:EU:C:2023:537), Randnummer 112; EuGH, Urteil vom 11. Dezember 2019, Rechtssache C-708/18, *Asociația de Proprietari bloc M5A-ScaraA* (ECLI:EU:C:2019:1064), Randnummer 58; EuGH, Urteil vom 4. Oktober 2024, Rechtssache C-621/22, *Koninklijke Nederlandse Lawn Tennisbond* (ECLI:EU:C:2024:857), Randnummer 55.
[73] So hat beispielsweise der EuGH im Urteil vom 4. Juli 2023, Rechtssache C-252/21, *Meta/Bundeskartellamt* (ECLI:EU:C:2023:537), Randnummer 123, festgestellt, dass die „Produktverbesserung“ nicht grundsätzlich als

Personen gegebenen Informationen berücksichtigt werden, um zu beurteilen, ob die betroffenen Personen vernünftigerweise mit der Verarbeitung ihrer personenbezogenen Daten rechnen konnten. Wenn keinerlei Informationen gegeben wurden, kann dies dafür sprechen, dass eine bestimmte Verarbeitung für die betroffenen Personen vernünftigerweise nicht zu erwarten war; werden die in der DSGVO vorgesehenen Transparenzanforderungen erfüllt, spricht dieser Umstand, für sich genommen, jedoch noch nicht dafür, dass eine bestimmte Verarbeitung für die betroffenen Personen vernünftigerweise zu erwarten war[74]. Nur weil die Datenschutzerklärung des Verantwortlichen Informationen enthält, die die Entwicklungsphase eines KI-Modells betreffen, bedeutet dies allein nicht notwendigerweise, dass die betroffenen Personen vernünftigerweise damit rechnen können, dass diese stattfindet; dies ist vielmehr von den Aufsichtsbehörden unter Berücksichtigung der besonderen Gegebenheiten des Einzelfalls und aller relevanten Umstände zu prüfen.

93. Bei der Analyse der vernünftigen Erwartungen der betroffenen Personen im Hinblick auf die in der Entwicklungsphase stattfindende Verarbeitung ist es wichtig, die Elemente zu berücksichtigen, die in den Leitlinien des EDSA über das berechtigte Interesse[75] genannt werden. Was den Gegenstand dieser Stellungnahme angeht, ist es darüber hinaus wichtig, den allgemeinen Kontext der Verarbeitung zu berücksichtigen. Dieser kann, wobei dies keine abschließende Aufzählung ist, Folgendes umfassen: ob die personenbezogenen Daten öffentlich verfügbar waren, die Art des Verhältnisses zwischen der betroffenen Person und dem Verantwortlichen (und ob zwischen ihnen eine Verbindung besteht), die Art der Dienstleistung, die Umstände, unter denen die personenbezogenen Daten erhoben wurden, die für die Datenerhebung genutzte Quelle (z. B. die Website oder der Dienst, die/der für die Erhebung der personenbezogenen Daten genutzt wurde, und die dort angebotenen Datenschutzeinstellungen), die potenziellen weiteren Verwendungen des Modells sowie ob sich die betroffenen Personen überhaupt dessen bewusst sind, dass ihre personenbezogenen Daten online sind.

94. In der Entwicklungsphase des Modells können die vernünftigen Erwartungen der betroffenen Personen differieren, je nachdem, ob die für die Modellentwicklung verarbeiteten Daten von den betroffenen Personen veröffentlicht wurden oder nicht. Die vernünftigen Erwartungen können auch unterschiedlich ausfallen, je nachdem, ob die betroffenen Personen die Daten direkt dem Verantwortlichen zur Verfügung gestellt haben (beispielsweise bei der Nutzung einer angebotenen Dienstleistung) oder ob der Verantwortliche die Daten aus einer anderen Quelle bezogen hat (z. B. von einem Dritten oder mittels Scraping). In beiden Fällen sollte für die Bewertung der vernünftigen Erwartungen auch betrachtet werden, was unternommen wurde, um die betroffenen Personen über die Verarbeitungstätigkeit zu informieren.

95. Gleichermaßen wichtig ist es, die vernünftigen Erwartungen der betroffenen Personen in der Einsatzphase des KI-Modells im Kontext der spezifischen Fähigkeiten des Modells zu betrachten. Bei KI-Modellen, die sich den Eingaben anpassen können, kann es zum Beispiel relevant sein, zu prüfen, ob sich die betroffenen Personen dessen bewusst waren, dass sie ihre personenbezogenen Daten mitgeteilt hatten, damit das KI-Modell seine Antworten ihren Bedürfnissen anpassen kann, sodass sie auf sie persönlich zugeschnittene Dienstleistungen erhalten können. Des Weiteren kann es auch

---

berechtigtes Interesse ausgeschlossen werden kann, wobei er jedoch im Weiteren ausgeführt hat, dass es *„in Anbetracht des Umfangs dieser Verarbeitung, ihrer erheblichen Auswirkungen auf den Nutzer sowie des Umstands, dass der Nutzer vernünftigerweise nicht damit rechnen kann, dass die Daten … verarbeitet werden, allerdings zweifelhaft [erscheint], ob das Ziel der Produktverbesserung Vorrang vor den Interessen und Grundrechten des Nutzers haben kann, zumal wenn es sich bei diesem um ein Kind handelt“*.

[74] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummer 53.

[75] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummern 50-54.

relevant sein, zu prüfen, ob diese Verarbeitungstätigkeit nur Auswirkungen auf die für die betroffenen Personen erbrachten Dienstleistungen hätte (z. B. bei der Personalisierung des Inhalts für einen bestimmten Nutzer) oder ob sie dazu genutzt würde, die Dienstleistung für sämtliche Kunden zu modifizieren (z. B. um das Modell allgemein zu verbessern). Wie bereits im Entwicklungsstadium kann es auch hier von besonderer Relevanz sein, zu prüfen, ob zwischen den betroffenen Personen und dem Verantwortlichen eine direkte Verbindung besteht. Im Falle einer solchen direkten Verbindung kann es dem Verantwortlichen beispielsweise möglich sein, die betroffenen Personen auf einfache Weise über die Verarbeitungstätigkeit und das Modell zu informieren, was wiederum Einfluss auf die vernünftigen Erwartungen der betroffenen Personen haben könnte.

Risikomindernde Maßnahmen

96. In Fällen, in denen die Interessen, Grundrechte und Grundfreiheiten der betroffenen Personen die vom Verantwortlichen oder einem Dritten verfolgten berechtigten Interessen überwiegen, kann der Verantwortliche in Betracht ziehen, risikomindernde Maßnahmen zu ergreifen, um die Auswirkungen der Verarbeitung auf die betroffenen Personen in Grenzen zu halten. Risikomindernde Maßnahmen sind Garantien, die auf die Umstände des Einzelfalls zugeschnitten sein sollten und von verschiedenen Umständen abhängig sind, unter anderem von der beabsichtigten Verwendung des KI-Modells. Solche risikomindernden Maßnahmen würden darauf abzielen, sicherzustellen, dass es keine Interessen gibt, die die Interessen des Verantwortlichen oder eines Dritten überwiegen, damit sich der Verantwortliche auf diese Rechtsgrundlage berufen könnte.

97. Allerdings hat der EDSA in seinen Leitlinien über berechtigte Interessen daran erinnert, dass risikomindernde Maßnahmen nicht mit denjenigen Maßnahmen verwechselt werden dürfen, zu denen der Verantwortliche – ganz unabhängig davon, ob die Verarbeitung auf Artikel 6 Absatz 1 Buchstabe f DSGVO[76] gestützt ist oder nicht – ohnehin gesetzlich verpflichtet ist, um die DSGVO-Konformität sicherzustellen. Dies gilt insbesondere für Maßnahmen zur Einhaltung der DSGVO-Grundsätze, beispielsweise des Grundsatzes der Datenminimierung.

98. Bei den nachstehend aufgeführten Maßnahmen handelt es sich um eine nicht erschöpfende und unverbindliche Auflistung und die Umsetzung dieser Maßnahmen sollte jeweils im Rahmen einer Einzelfallbewertung in Betracht gezogen werden. Je nach den Umständen des Einzelfalls kann es sein, dass einige der nachstehenden Maßnahmen zur Einhaltung bestimmter Pflichten aus der DSGVO erforderlich sind; ist dies nicht der Fall, so können sie als zusätzliche Garantien in Betracht gezogen werden. Zudem beziehen sich einige der nachstehend genannten Maßnahmen auf Bereiche, die rapider Weiterentwicklung und Neuentwicklungen unterliegen, und dies sollte von den Aufsichtsbehörden bei der Bearbeitung des Einzelfalls berücksichtigt werden.

99. **In Bezug auf die Entwicklungsphase der KI-Modelle** können verschiedene Vorkehrungen getroffen werden, um die Risiken zu mindern, die sich ergeben, wenn vom Verantwortlichen selbst erhobene bzw. von Dritten erhobene Daten verarbeitet werden (wozu auch die Minderung der sich aus Web Scraping ergebenden Risiken gehört). Vor diesem Hintergrund zeigt der EDSA einige Beispiele für Maßnahmen auf, mit denen in der Abwägungsprüfung erkannte Risiken gemindert werden können; bei der Bewertung spezifischer KI-Modelle sollten die Aufsichtsbehörden diese Maßnahmen auf Einzelfallbasis in Betracht ziehen.

100. **Technische Maßnahmen:**

[76] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummer 57.

a. In Abschnitt 3.2.2 genannte Maßnahmen, die geeignet sind, die in Rede stehenden Risiken zu mindern, wobei diese Maßnahmen nicht zur Anonymisierung des Modells führen und weder zur Einhaltung anderer DSGVO-Pflichten noch gemäß der Erforderlichkeitsprüfung (dem zweiten Schritt der Prüfung der berechtigten Interessen) erforderlich sind.

101. Andere relevante Maßnahmen, die darüber hinaus in Betracht kommen, sind beispielsweise:

b. Pseudonymisierungsmaßnahmen: hierunter könnten etwa Maßnahmen fallen, die verhindern, dass Daten auf Grundlage einzelner Identifikatoren kombiniert werden. Diese Maßnahmen sind unter Umständen nicht geeignet, wenn die Aufsichtsbehörde befindet, dass der Verantwortliche Nachweis dafür erbracht hat, dass es angemessenerweise erforderlich war, für die Entwicklung des betreffenden KI-Systems oder -Modells andere Daten über eine bestimmte natürliche Person zu sammeln.

c. Maßnahmen, mit denen personenbezogene Daten im Trainingsdatensatz maskiert oder durch fingierte personenbezogene Daten ersetzt werden (z. B. Ersetzung der Namen und E-Mail-Adressen durch Platzhalter-Namen und -E-Mail-Adressen). Diese Maßnahme kommt insbesondere in Betracht, wenn der tatsächliche materielle Gehalt der Daten für die Gesamtverarbeitung nicht relevant ist (z. B. beim Training von großen Sprachmodellen).

102. **Maßnahmen zur Ermöglichung der Rechtsausübung durch natürliche Personen:**

a. Einhaltung einer angemessenen Frist zwischen der Erhebung des Trainingsdatensatzes und seiner Verwendung. Diese zusätzliche Garantie kann den betroffenen Personen ermöglichen, ihre Rechte auf Löschung auszuüben, wobei die Angemessenheit der Frist nach den Umständen des Einzelfalls zu beurteilen ist.

b. Einräumung der Möglichkeit, gleich zu Beginn den uneingeschränkten „Opt-out“ zu erklären, beispielsweise indem es ins Ermessen der betroffenen Personen gestellt wird, den Widerspruch zu erklären, bevor die Verarbeitung beginnt; auf diese Weise wird den Personen eine stärkere Kontrolle über ihre Daten eingeräumt, die noch über das Widerspruchsrecht in Artikel 21 DSGVO[77] hinausgeht.

c. Einräumung der Möglichkeit, dass betroffene Person ihr Recht auf Löschung auch dann ausüben können, wenn keiner der in Artikel 17 Absatz 1 DSGVO aufgeführten Gründe vorliegt[78].

d. Einräumung der Möglichkeit, dass betroffene Personen Ansprüche wegen Regurgitation oder Memorisation personenbezogener Daten geltend machen sowie die Umstände und Mittel mitteilen können, unter bzw. mit denen die geltend gemachten Vorfälle reproduziert werden können, damit die Verantwortlichen diese reproduzieren und einschlägige relevante Techniken für das Entlernen bewerten können, um die Vorfälle abzustellen.

103. **Transparenzmaßnahmen**: Zuweilen könnten risikomindernde Maßnahmen auch Maßnahmen beinhalten, die mehr Transparenz bezüglich der Entwicklung des KI-Modells bieten. Es gibt Maßnahmen, die zusätzlich zur Einhaltung der DSGVO-Pflichten dazu beitragen können, die Informationsasymmetrie zu überwinden und betroffenen Personen zu ermöglichen, die in der Entwicklungsphase stattfindende Verarbeitung besser zu verstehen:

[77] Ebenda.
[78] Ebenda.

a. Herausgabe öffentlicher und leicht zugänglicher Mitteilungen, die über die Informationspflichten gemäß den Artikeln 13 und 14 DSGVO hinausgehen; beispielsweise durch zusätzliche Angaben über die Erhebungskriterien und alle verwendeten Datensätze, unter Berücksichtigung besonderer Schutzvorkehrungen für Kinder und vulnerable Personen.

b. Alternative Formen der Information der betroffenen Personen, zum Beispiel: Kampagnen in verschiedenen Medien zur Information der betroffenen Personen, E-Mail-Informationskampagnen, Nutzung grafischer Visualisierung, häufig gestellte Fragen (FAQ), Transparenzsiegel und Modellkarten, deren Systematik die Darstellung der Informationen über KI-Modelle strukturieren könnte, sowie jährliche Transparenzberichte auf freiwilliger Basis.

104. **Spezifische risikomindernde Maßnahmen im Zusammenhang mit Web Scraping:** Da, wie vorstehend erwähnt, das Web Scraping spezifische Risiken birgt[79], könnten in diesem Zusammenhang spezifische risikomindernde Maßnahmen ermittelt werden. Soweit relevant, könnten diese von den Aufsichtsbehörden zusätzlich zu den vorgenannten risikomindernden Maßnahmen in Betracht gezogen werden, wenn die von ihnen untersuchten Verantwortlichen Web Scraping durchführen.

105. Spezifische Maßnahmen könnten sich, selbst wenn sie im zweiten Schritt der Prüfung der berechtigten Interessen nicht für erforderlich befunden wurden, durchaus als nützlich für die Minderung der im Zusammenhang mit Web Scraping bestehenden Risiken erweisen. Dazu können **technische Maßnahmen** gehören, beispielsweise:

a. der Ausschluss von Dateninhalten aus Veröffentlichungen, die personenbezogene Daten enthalten können, die Risiken für bestimmte Personen oder Personengruppen mit sich bringen könnten (z. B. Personen, die, falls die Informationen öffentlich freigegeben würden, Beleidigungen, Vorurteilen oder sogar körperlicher Gewalt ausgesetzt sein könnten).

b. die Sicherstellung, dass gewisse Datenkategorien nicht erhoben oder dass bestimmte Quellen von der Datenerhebung ausgeschlossen werden; das könnte beispielsweise bestimmte Websites betreffen, bei denen wegen der Empfindlichkeit ihres Gegenstands der Eingriff in die Privatsphäre besonders groß ist.

c. der Ausschluss der Datenerhebung von Websites oder (Abschnitten von Websites), die klar dem Web Scraping und der Wiederverwendung ihres Inhalts für den Aufbau von Datenbanken für das KI-Training widersprochen haben (beispielsweise durch Beachtung der robots.txt- oder ai.txt-Dateien oder andere anerkannte Mechanismen, die den Ausschluss von automatisiertem Crawling oder Scraping zum Ausdruck bringen).

d. die Auferlegung anderer relevanter Einschränkungen der Datenerhebung, möglicherweise einschließlich fristbezogener Kriterien.

106. Beispiele für spezifische Maßnahmen im Zusammenhang mit Web Scraping, die den **natürlichen Personen die Rechtsausübung ermöglichen und für Transparenz sorgen**, sind: die Einrichtung einer vom Verantwortlichen verwalteten Opt-out-Liste, die den betroffenen Personen schon vor Beginn der Datenerhebung die Möglichkeit gibt, der Erhebung ihrer Daten auf bestimmten Websites oder Online-

[79] Diese Praktiken könnten weitere Probleme aufwerfen, die nicht Gegenstand dieser Stellungnahme sind, vgl. zum Beispiel Pagallo U., Ciani Sciolla J., *Anatomy of web data scraping: ethics, standards, and the troubles of the law.* European Journal of Privacy Law & Technologies, 2023, Ausgabe 2, S. 1-19, abrufbar unter: https://doi.org/10.57230/EJPLT232PS.

Plattformen zu widersprechen, indem sie Informationen zur Verfügung stellen, die sie auf diesen Websites identifizieren[80].

107. **<u>Spezifische Erwägungen in Bezug auf risikomindernde Maßnahmen in der Einsatzphase</u>:** Je nach den Umständen, mögen einige der vorgenannten Maßnahmen auch für die Einsatzphase relevant sein; nachstehend bietet der EDSA jedoch eine nicht erschöpfende Liste zusätzlicher Unterstützungsmaßnahmen, die ergriffen werden könnten und die von den Aufsichtsbehörden auf Einzelfallbasis zu bewerten sind.

    a. **Technische Maßnahmen** können beispielsweise ergriffen werden, um – insbesondere im Zusammenhang mit generativen KI-Modellen – die Speicherung, Regurgitation oder Erzeugung personenbezogener Daten zu verhindern (zum Beispiel durch Ausgabefilter), und/oder um das Risiko der rechtswidrigen Wiederverwendung durch KI-Modelle mit allgemeinem Verwendungszweck zu mindern (z. B. durch digitale Wasserzeichen für KI-generierte Ausgaben).

    b. **Maßnahmen zur Erleichterung oder Beschleunigung der Rechtsausübung durch natürliche Personen** in der Einsatzphase, die über die gesetzlich vorgeschriebenen Maßnahmen hinausgehen – insbesondere, wobei dies keine abschließende Aufzählung ist, in Bezug auf die Ausübung des Rechts auf Löschung personenbezogener Daten aus Modellausgabedaten oder Deduplizierung – sowie nach der Trainingsphase (Post-Training) eingesetzte Techniken, die personenbezogene Daten löschen oder unterdrücken sollen.

108. Bei der Untersuchung des Einsatzes spezifischer KI-Modelle sollten die Aufsichtsbehörden berücksichtigen, ob der Verantwortliche die von ihm durchgeführte Abwägungsprüfung veröffentlicht hat, da ein solches Vorgehen die Transparenz und Fairness erhöhen könnte. Wie bereits in den Leitlinien des EDSA über berechtigte Interessen erwähnt wurde, kommen auch andere Maßnahmen in Betracht, um die betroffenen Personen bereits vor der Erhebung ihrer personenbezogenen Daten über die Abwägungsprüfung zu informieren[81]. Der EDSA erinnert auch nochmals[82] daran, das auch zu berücksichtigen ist, ob gegebenenfalls der DSB hinzugezogen wurde.

3.4 Mögliche Auswirkungen einer rechtswidrigen Verarbeitung in der Entwicklung des KI-Modells auf die Rechtmäßigkeit der späteren Verarbeitung oder des späteren Betriebs des KI-Modells

109. Dieser Abschnitt der Stellungnahme betrifft Frage 4 des Ersuchens. Mit dieser Frage wird um Klarstellung ersucht, welche Auswirkungen eine rechtswidrige Verarbeitung in der Entwicklungsphase auf die Rechtmäßigkeit der späteren Verarbeitung (etwa in der Einsatzphase des KI-Modells) oder auf den späteren Betrieb des KI-Modells haben kann. Die Frage zielt sowohl auf die Situation ab, dass ein KI-Modell im Modell enthaltene personenbezogene Daten verarbeitet (Frage 4(i) des Ersuchens), als auch auf die Situation, dass beim Einsatz des KI-Modells keine personenbezogenen Daten mehr verarbeitet werden (d. h. das Modell ist anonym) (Frage 4(ii) des Ersuchens).

[80] Es sei denn, der Verantwortliche kann zwingende schutzwürdige Gründe für die Verarbeitung nachweisen, die die Interessen, Rechte und Freiheiten der betroffenen Person überwiegen, oder die Verarbeitung dient der Geltendmachung, Ausübung oder Verteidigung von Rechtsansprüchen.
[81] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummer 68.
[82] Vgl. EDSA, Leitlinien 1/2024 zur Verarbeitung personenbezogener Daten auf der Grundlage von Artikel 6 Absatz 1 Buchstabe f DSGVO, Version 1.0, angenommen am 8. Oktober 2024, Randnummer 12.

110. Bevor er auf spezifische Szenarien eingeht, schickt der EDSA folgende allgemeine Erwägungen voraus.

111. Erstens wird der Fokus der in diesem Abschnitt gegebenen Klarstellungen darauf liegen, die Verarbeitung personenbezogener Daten zu untersuchen, die in der Entwicklungsphase unter Verstoß gegen den Grundsatz der Rechtmäßigkeit in Artikel 5 Absatz 1 Buchstabe a DSGVO und insbesondere Artikel 6 DSGVO erfolgt (im Folgenden „**Rechtswidrigkeit**")[83]. In diesem Sinne werden die Erwägungen des EDSA darauf fokussieren, wie sich die Rechtswidrigkeit der Verarbeitung in der Entwicklungsphase auf die Rechtmäßigkeit der anschließenden Verarbeitung oder des späteren Betriebs des Modells auswirkt (es geht also um die Einhaltung von Artikel 5 Absatz 1 Buchstabe a DSGVO und Artikel 6 DSGVO). Der EDSA weist jedoch darauf hin, dass die in der Entwicklungsphase durchgeführte Verarbeitung auch zu Verstößen gegen andere DSGVO-Vorschriften führen kann, die nicht Gegenstand dieser Stellungnahme sind; etwa zu Verstößen wegen mangelnder Transparenz gegenüber betroffenen Personen oder fehlenden Datenschutzes durch Technikgestaltung und datenschutzfreundliche Voreinstellungen.

112. Zweitens ist für die Beantwortung dieser Frage der Grundsatz der Rechenschaftspflicht entscheidend, wonach Verantwortliche dafür verantwortlich sind, unter anderem Artikel 5 Absatz 1 DSGVO und Artikel 6 DSGVO einzuhalten[84] und dafür Nachweis zu erbringen. Dies gilt auch für das Erfordernis, zu prüfen, welche Organisation der für die betreffende Verarbeitungstätigkeit Verantwortliche ist und ob sich (wegen untrennbarer Verbundenheit) Situationen mit gemeinsamer Verantwortlichkeit ergeben[85]. Da es stets auf die Umstände des Einzelfalls – auch auf die Rollen jede der an der Verarbeitung beteiligten Parteien – ankommt, sind die Ausführungen des EDSA als allgemeine Bemerkungen zu verstehen, die von den Aufsichtsbehörden jeweils auf Einzelfallbasis zu bewerten sind.

113. Drittens hebt der EDSA hervor, dass gemäß Artikel 51 Absatz 1 DSGVO die Aufsichtsbehörden *„für die Überwachung der Anwendung [der DSGVO] zuständig sind, damit die Grundrechte und Grundfreiheiten natürlicher Personen bei der Verarbeitung geschützt werden und der freie Verkehr personenbezogener Daten in der Union erleichtert wird"*. Es liegt deshalb in der Zuständigkeit der Aufsichtsbehörden, die Rechtmäßigkeit der Verarbeitung zu bewerten und die ihr durch die DSGVO eingeräumten Befugnisse im Einklang mit ihrem nationalen Rahmenwerk[86] auszuüben. In derartigen Fällen haben die Aufsichtsbehörden einen Ermessensspielraum für die Bewertung etwaiger Datenschutzverletzungen sowie, soweit erforderlich, für die Auswahl aus den in Artikel 58 DSGVO aufgeführten geeigneten, erforderlichen und verhältnismäßigen Maßnahmen, wobei die Umstände des Einzelfalls zu berücksichtigen sind[87].

---

[83] EuGH, Urteil vom 4. Mai 2023, Rechtssache C-60/22, *Bundesrepublik Deutschland* (ECLI:EU:C:2023:373), Randnummern 55-57.

[84] EuGH, Urteil vom 4. Mai 2023, Rechtssache C-60/22, *Bundesrepublik Deutschland* (ECLI:EU:C:2023:373), Randnummern 53.

[85] Leitlinien 07/2020 des EDSA zu den Begriffen „Verantwortlicher" und „Auftragsverarbeiter" in der DSGVO, Version 2.1, angenommen am 7. Juli 2021, Randnummer 55.

[86] Unter Umständen sind einschlägige nationale Vorschriften zu beachten. Vgl. zum Beispiel Artikel 2-decies des italienischen Datenschutzgesetzes (Decreto Legislativo Nr. 196/2003), der bestimmt, dass Daten, die unter Verstoß gegen die Datenschutzvorschriften verarbeitet wurden, nicht verwendet werden dürfen. Andere nationale Gesetzeswerke, beispielsweise Strafvorschriften, bleiben unberührt.

[87] Vgl. insoweit Erwägungsgrund 129 der DSGVO, wie auch EuGH, Urteil vom 26. September 2024, Rechtssache C-768/21, *TR / Land Hessen* (ECLI:EU:C:2024:785), Randnummer 37; Urteil des EuGH vom 7. Dezember 2023, verbundene Rechtssachen C-26/22 und C-64/22, SCHUFA Holding (Libération de reliquat de dette) (ECLI:EU:C:2023:958), Randnummer 57; sowie EuGH, Urteil vom 14. März 2024, Rechtssache C-46/23, *Újpesti Polgármesteri Hivatal* (ECLI:EU:C:2024:239), Randnummer 34.

114. **Wird eine Datenschutzverletzung festgestellt, so können die Aufsichtsbehörden Abhilfemaßnahmen anordnen, indem sie beispielsweise – unter Berücksichtigung der Umstände des Einzelfalls – die Verantwortlichen anweisen, Maßnahmen zu ergreifen, die die Rechtswidrigkeit der anfänglichen Verarbeitung beheben.** In Betracht kommt beispielsweise, eine Geldbuße zu verhängen, die Verarbeitung vorübergehend zu beschränken, einen Teil des rechtswidrig verarbeiteten Datensatzes zu löschen oder, falls das nicht möglich sein sollte, je nach den Gegebenheiten des Sachverhalts und unter Berücksichtigung der Verhältnismäßigkeit der Maßnahme die Löschung des gesamten für die Entwicklung des KI-Modells verwendeten Datensatzes und/oder des eigentlichen KI-Modells anzuordnen. Bei der Bewertung der Verhältnismäßigkeit der vorgesehenen Maßnahme können die Aufsichtsbehörden Maßnahmen, die der Verantwortliche zur Behebung der Rechtswidrigkeit der anfänglichen Verarbeitung ergreifen kann (z. B. erneutes Training (Retraining)), berücksichtigen.

115. Des Weiteren weist der EDSA darauf hin, dass betroffene Personen, deren personenbezogene Daten rechtswidrig verarbeitet werden, unter den in Artikel 17 DSGVO genannten Voraussetzungen die Löschung ihrer personenbezogenen Daten verlangen können und dass die Aufsichtsbehörden die Löschung der personenbezogenen Daten von Amts wegen anordnen können[88].

116. Für die Bewertung, ob eine Maßnahme geeignet, erforderlich und verhältnismäßig ist, können die Aufsichtsbehörden unter anderem auch die für die betroffenen Personen entstandenen Risiken, den Schweregrad der Datenschutzverletzung, die technische und finanzielle Umsetzbarkeit der Maßnahme sowie die Menge der betroffenen personenbezogenen Daten berücksichtigen.

117. Abschließend erinnert der EDSA daran, dass die von den Aufsichtsbehörden gemäß der DSGVO ergriffenen Maßnahmen die Maßnahmen unberührt lassen, die von zuständigen Behörden gemäß der KI-Verordnung und/oder anderen einschlägigen Rechtsrahmen (z. B. Gesetze über zivilrechtliche Haftung) getroffen werden.

118. In den anschließenden Abschnitten wird der EDSA die drei unter Frage 4 des Ersuchens fallenden Szenarien behandeln, welche sich jeweils darin unterscheiden, ob die für die Modellentwicklung verarbeiteten personenbezogenen Daten im Modell verbleiben und/oder ob die spätere Verarbeitung von demselben oder von einem anderen Verantwortlichen durchgeführt wird.

3.4.1 Szenario 1: Für die Modellentwicklung verarbeitet der Verantwortliche rechtswidrig personenbezogene Daten, die personenbezogenen Daten verbleiben im Modell und werden später von demselben Verantwortlichen verarbeitet (z. B. im Zusammenhang mit dem Einsatz des Modells)

119. Dieses Szenario betrifft Frage 4(i) des Ersuchens, nämlich die Situation, dass der Verantwortliche rechtswidrig (d. h. unter Verstoß gegen Artikel 5 Absatz 1 Buchstabe a DSGVO und Artikel 6 DSGVO) personenbezogene Daten verarbeitet, um ein KI-Modell zu entwickeln, wobei das KI-Modell nicht anonym ist, weil es Informationen enthält, die eine identifizierte oder identifizierbare natürliche Person betreffen. Die personenbezogenen Daten werden später von demselben Verantwortlichen

[88] Vgl. dazu *EDPB Opinion 39/2021 on whether Article 58(2)(g) GDPR could serve as a legal basis for a supervisory authority to order ex officio the erasure of personal data in a situation where such request was not submitted by the data subject* (EDSA, Stellungnahme 39/2021 zu der Frage, ob Artikel 58 Absatz 2 Buchstabe g DSGVO als Rechtsgrundlage für die aufsichtsbehördliche Anordnung der Löschung personenbezogener Daten dienen könnte, wenn dies nicht von der betroffenen Person verlangt wurde), Randnummer 28. Vgl. dazu auch EuGH, Urteil vom 14. März 2024, Rechtssache C-46/23, *Újpesti Polgármesteri Hivatal* (ECLI:EU:C:2024:239), Randnummer 42.

verarbeitet (z. B. im Zusammenhang mit dem Einsatz des Modells). Zu diesem Szenario stellt der EDSA folgende Erwägungen an:

120. Die Befugnis der Aufsichtsbehörde zur Anordnung von Abhilfemaßnahmen in Bezug auf die erste Verarbeitung (vgl. dazu die vorstehenden Erklärungen in den Randnummern 113, 114, 115) hätte grundsätzlich Auswirkungen auf die spätere Verarbeitung (sollte z. B. die Aufsichtsbehörde anordnen, dass der Verantwortliche die rechtswidrig verarbeiteten personenbezogenen Daten löscht, so würde diese Abhilfemaßnahme nicht zulassen, dass Letzterer die personenbezogenen Daten, denen die Maßnahmen galten, später verarbeitet).

121. Spezifisch im Hinblick auf die Auswirkungen der rechtswidrigen Verarbeitung in der Entwicklungsphase auf die spätere Verarbeitung (z. B. in der Einsatzphase) erinnert der EDSA daran, dass die Aufsichtsbehörden gehalten sind, eine Einzelfallprüfung vorzunehmen, die die besonderen Umstände des betreffenden Falls berücksichtigt.

122. **Ob in der Entwicklungs- und Einsatzphase jeweils gesonderte Zwecke verfolgt werden (sodass es sich um gesonderte Verarbeitungsvorgänge handelt) und inwieweit das Fehlen einer Rechtsgrundlage für die anfängliche Verarbeitungstätigkeit Auswirkungen auf die Rechtmäßigkeit der späteren Verarbeitung hat, sollte auf Einzelfallbasis je nach den Umständen des Einzelfalls bewertet werden.**

123. Was insbesondere die Rechtsgrundlage in Artikel 6 Absatz 1 Buchstabe f DSGVO angeht, so sollte, wenn die spätere Verarbeitung auf berechtigte Interessen gestützt ist, der Umstand, dass die erste Verarbeitung rechtswidrig war, bei der Prüfung der berechtigten Interessen berücksichtigt werden (z. B. im Hinblick auf die Risiken für die betroffenen Personen oder auf den Umstand, dass die betroffenen Personen möglicherweise nicht mit einer solchen späteren Verarbeitung rechnen). In diesen Fällen kann die Rechtswidrigkeit der Verarbeitung in der Entwicklungsphase Auswirkungen auf die Rechtmäßigkeit der späteren Verarbeitung haben.

3.4.2 Szenario 2: Für die Modellentwicklung verarbeitet der Verantwortliche rechtswidrig personenbezogene Daten, die personenbezogenen Daten verbleiben im Modell und werden später im Zusammenhang mit dem Einsatz des Modells von einem anderen Verantwortlichen verarbeitet

124. Dieses Szenario betrifft Frage 4(i) des Ersuchens. Von Szenario 1 (in Abschnitt 3.4.1 dieser Stellungnahme) unterscheidet es sich insofern, als die personenbezogenen Daten später im Zusammenhang mit dem Einsatz des KI-Modells von einem anderen Verantwortlichen verarbeitet werden.

125. Der EDSA erinnert daran, dass die Prüfung der Rollen, die diesen verschiedenen Akteuren nach dem Datenschutzregelwerk zugewiesen sind, unerlässliche Voraussetzung dafür ist, die einschlägigen DSGVO-Pflichten sowie die persönliche Verantwortlichkeit für diese Pflichten zu ermitteln; dabei sollten für die Bewertung der sich aus der DSGVO ergebenden Verantwortlichkeiten der Parteien auch Situationen mit gemeinsamer Verantwortlichkeit in Betracht gezogen werden. Die nachstehenden Bemerkungen sind deshalb als allgemeine Erwägungen anzusehen, die gegebenenfalls von den Aufsichtsbehörden berücksichtigt werden sollten. Zu diesem Szenario 2 stellt der EDSA folgende Erwägungen an:

126. Erstens ist daran zu erinnern, dass gemäß Artikel 5 Absatz 1 Buchstabe a DSGVO in Verbindung mit Artikel 5 Absatz 2 DSGVO jeder Verantwortliche die Rechtmäßigkeit der von ihm durchgeführten Verarbeitung sicherstellen und dafür Nachweis erbringen können muss. Die Aufsichtsbehörden sollten die Rechtmäßigkeit der Verarbeitungen bewerten, und zwar sowohl (i) die Verarbeitungen des

Verantwortlichen, von dem das KI-Modell ursprünglich entwickelt wurde, als auch (ii) die Verarbeitungen des Verantwortlichen, der das KI-Modell erworben hat und nun die personenbezogenen Daten selbst verarbeitet.

127. Zweitens sind die obigen Ausführungen in den Randnummern 113, 114, 115 zur aufsichtsbehördlichen Eingriffsbefugnis in Bezug auf die anfängliche Verarbeitung auch in diesem Fall relevant. Je nach den Umständen des Einzelfalls können in diesem Zusammenhang auch Artikel 17 Absatz 1 Buchstabe d DSGVO (Löschung rechtswidrig verarbeiteter Daten) und Artikel 19 DSGVO (Mitteilungspflicht im Zusammenhang mit der Berichtigung oder Löschung personenbezogener Daten oder der Einschränkung der Verarbeitung) relevant sein, beispielsweise im Hinblick auf die Mitteilungspflicht des Verantwortlichen, von dem das Modell entwickelt wurde, gegenüber dem das Modell einsetzenden Verantwortlichen.

128. Drittens sollte diese Bewertung der etwaigen Auswirkungen der Rechtswidrigkeit der anfänglichen Verarbeitung auf die spätere, von einem anderen Verantwortlichen durchgeführte Verarbeitung von den Aufsichtsbehörden auf Einzelfallbasis vorgenommen werden.

129. **Die Aufsichtsbehörden sollten berücksichtigen, ob sich der Verantwortliche, der das Modell einsetzt, im Rahmen seiner Rechenschaftspflichten[89] zum Nachweis der Einhaltung von Artikel 5 Absatz 1 Buchstabe a DSGVO und Artikel 6 DSGVO in geeigneter Weise vergewissert hat, dass für die Entwicklung des KI-Modells keine rechtswidrig verarbeiteten personenbezogenen Daten verarbeitet wurden.** Die Aufsichtsbehörden sollten bei der Bewertung berücksichtigen, ob der Verantwortliche in der nicht erschöpfenden Aufzählung aufgeführte Kriterien – zum Beispiel die Datenquelle bzw. ob das KI-Modell unter Verstoß gegen die DSGVO erstellt wurde – geprüft hat; dies gilt insbesondere, wenn derartige Verstöße durch eine Aufsichtsbehörde oder ein Gericht festgestellt wurden, sodass die Rechtswidrigkeit der anfänglichen Verarbeitung für den das Modell einsetzenden Verantwortliche nicht zu ignorieren war.

130. Der Verantwortliche sollte beispielsweise berücksichtigen, ob die Daten aus einer Verletzung des Schutzes personenbezogener Daten stammen oder ob die Verarbeitung Gegenstand einer aufsichtsbehördlichen oder gerichtlichen Verstoßfeststellung war. **Die Erwartungen der Aufsichtsbehörden an die Gründlichkeit und Detailliertheit der vom Verantwortlichen vorzunehmenden Bewertung kann von verschiedenen Faktoren abhängen, unter anderem von der Art und dem Grad der Risiken, die sich durch die beim Einsatz des KI-Modells erfolgende Verarbeitung für die betroffenen Personen, deren Daten für die Modellentwicklung verwendet wurden, ergeben.**

131. Der EDSA merkt an, dass Anbieter von Hochrisiko-KI-Systemen nach der KI-Verordnung verpflichtet sind, eine EU-Konformitätserklärung auszustellen[90], mit der erklärt wird, dass das maßgebliche KI-System dem EU-Datenschutzrecht entspricht[91]. Der EDSA merkt an, dass eine solche Eigenerklärung nicht notwendigerweise den Schluss auf DSGVO-Konformität zulässt. Sie kann aber dennoch bei der Untersuchung eines bestimmten KI-Modells von den Aufsichtsbehörden berücksichtigt werden.

132. Die obigen Erwägungen unter Randnummer 123 sind auch in diesem Fall relevant. Wenn Aufsichtsbehörden überprüfen, ob und in welcher Weise der Verantwortliche die Angemessenheit der berechtigten Interessen als Rechtsgrundlage für die von ihm durchgeführte Verarbeitung bewertet hat, sollte die Rechtswidrigkeit der anfänglichen Verarbeitung bei der Prüfung der berechtigten

[89] Artikel 5 Absatz 2 DSGVO und Artikel 24 DSGVO.
[90] Artikel 16 Buchstabe g und Artikel 47 der KI-Verordnung.
[91] Anhang V, Nummer 5 KI-Verordnung.

Interessen berücksichtigt werden; beispielsweise bei der Bewertung der potenziellen Risiken, die sich für die betroffenen Personen ergeben können, deren personenbezogene Daten rechtswidrig für die Modellentwicklung verarbeitet wurden. Im Rahmen der Abwägungsprüfung sind verschiedene Aspekte zu berücksichtigen, und zwar solche technischer Art (z. B. Vorhandensein bei der Modellentwicklung gesetzter Filter oder Zugangsbeschränkungen, die der spätere Verantwortliche weder umgehen noch beeinflussen kann und die den Zugriff auf personenbezogene Daten oder deren Offenlegung verhindern könnten) wie auch solche rechtlicher Art (z. B. Art und Schweregrad der Rechtswidrigkeit der anfänglichen Verarbeitung).

3.4.3 Szenario 3 Ein Verantwortlicher verarbeitet personenbezogene Daten in rechtswidriger Weise, um das Modell zu entwickeln; anschließend stellt er sicher, dass das Modell anonymisiert wird, bevor er selbst oder ein anderer Verantwortlicher im Zuge des Einsatzes eine weitere Verarbeitung personenbezogener Daten einleitet.

133. Dieses Szenario bezieht sich auf Frage 4(ii) des Ersuchens und betrifft den Fall, dass ein Verantwortlicher rechtswidrig personenbezogene Daten verarbeitet, um ein KI-Modell zu entwickeln, dabei jedoch so vorgeht, dass sichergestellt ist, dass die personenbezogenen Daten anonymisiert werden, bevor der Verantwortliche selbst oder ein anderer Verantwortlicher im Zuge des Modelleinsatzes eine weitere Verarbeitung personenbezogener Daten einleitet. Zunächst erinnert der EDSA daran, dass die Zuständigkeit und Eingriffsbefugnis der Aufsichtsbehörden sowohl für die Verarbeitung für die Zwecke der Anonymisierung des Modells als auch für die in der Entwicklungsphase durchgeführte Verarbeitung gilt. Die Aufsichtsbehörden können also je nach den Umständen des Einzelfalls Abhilfemaßnahmen hinsichtlich der anfänglichen Verarbeitung anordnen (siehe dazu die Erläuterungen oben in den Randnummern 113, 114 und 115).

134. Sofern Nachweis dafür erbracht werden kann, dass der spätere Betrieb des KI-Modells nicht mit der Verarbeitung personenbezogener Daten einhergeht, ist die DSGVO nach Ansicht des EDSA nicht anwendbar[92]. Die Rechtswidrigkeit der anfänglichen Verarbeitung sollte folglich keine Auswirkungen auf den späteren Betrieb des Modells haben. Der EDSA hebt jedoch hervor, dass eine bloß behauptete Anonymität des Modells nicht genügt, dieses von der Anwendung der DSGVO auszunehmen; die Aufsichtsbehörden sollten eine Einzelfallbewertung vornehmen, wobei die vom EDSA zur Beantwortung der Frage 1 des Ersuchens angestellten Erwägungen berücksichtigt werden sollten.

135. **Wenn die Verantwortlichen dann nach der Anonymisierung des Modells personenbezogene Daten verarbeiten, die erst in der Einsatzphase erhoben wurden, wäre die DSGVO auf diese Verarbeitungsvorgänge anwendbar. In diesen Fällen sollte, was die DSGVO angeht, die Rechtswidrigkeit der anfänglichen Datenverarbeitung die Rechtmäßigkeit der späteren, in der Einsatzphase ausgeführten Verarbeitung unberührt lassen.**

[92] Erwägungsgrund 26 der DSGVO.

4 Abschließende Bemerkungen

136. Diese Stellungnahme richtet sich an alle Aufsichtsbehörden und wird gemäß Artikel 64 Absatz 5 Buchstabe b DSGVO veröffentlicht.

Für den Europäischen Datenschutzausschuss

Die Vorsitzende

Anu Talus