DeutscherAnwaltVerein

Stellungnahme

des Deutschen Anwaltvereins vorbereitet durch den Ausschuss Informationsrecht

zu den Bestimmungen über künstliche Intelligenz im Digitalomnibuspaket (Kommissionvorschläge COM(2025) 837 final und COM(2025) 836 final)

Stellungnahme Nr.: 24/2026 Brüssel, im März 2026

Mitglieder des Ausschusses

- Rechtsanwalt Prof Niko Härting, Berlin (Vorsitzender)
- Rechtsanwalt Dr. Simon Assion, Frankfurt am Main
- Rechtsanwältin Dr. Christiane Bierekoven, Düsseldorf
- Rechtsanwältin Isabell Conrad, München (Berichterstatterin)
- Rechtsanwalt Prof. Dr. Malte Grützmacher, LL.M., Hamburg (Berichterstatter)
- Rechtsanwalt Peter Huppertz, LL.M, Düsseldorf
- Rechtsanwalt Dr. Helmut Redeker, Bonn
- Rechtsanwältin Dr. Kristina Schreiber, Köln
- Rechtsanwalt Dr. Robert Selk, LL.M. (EU), München

Zuständig in der DAV-Geschäftsstelle

- Rechtsanwältin Nicole Narewski, Geschäftsführerin, Berlin

Ansprechpartner in Brüssel

- Rechtsanwältin Dorothee Wildt, LL.M., stellv. Leiterin DAV Brüssel
- Myra Jockisch, LL.M., Referentin

Deutscher Anwaltverein
Littenstraße 11, 10179 Berlin
Tel.: +49 30 726152-0
Fax: +49 30 726152-190
E-Mail: dav@anwaltverein.de

Büro Brüssel
Rue Joseph II 40, Boîte 7B
1000 Brüssel, Belgien
Tel.: +32 2 28028-12
Fax: +32 2 28028-13
E-Mail: bruessel@eu.anwaltverein.de
EU-Transparenz-Registernummer:
87980341522-66

www.anwaltverein.de

Der Deutsche Anwaltverein (DAV) ist der freiwillige Zusammenschluss der deutschen Rechtsanwältinnen und Rechtsanwälte. Der DAV versammelt mehr als 60.000 Rechtsanwältinnen und Rechtsanwälte sowie Anwaltsnotarinnen und Anwaltsnotare, die in 253 lokalen Anwaltvereinen im In- und Ausland organisiert sind. Er vertritt die Interessen der deutschen Anwaltschaft auf nationaler, europäischer und internationaler Ebene. Der DAV ist im Lobbyregister für die Interessenvertretung gegenüber dem Deutschen Bundestag und der Bundesregierung zur Registernummer R000952 eingetragen.

Der Deutsche Anwaltverein nimmt nachfolgend Stellung zu den Entwürfen der Änderungsverordnungen der EU-Kommission COM (2025) 837 final (Teil 1) sowie COM(2025) 836 final (Teil 2) in ihren Fassungen vom 19.11.2025.

Teil 1: Änderungsvorschläge in Bezug auf die Datenschutzgrundverordnung

Im Grundsatz begrüßt der DAV den Versuch, durch Änderungen des Art. 9 DSGVO und einen neuen Art. 88c DSGVO die spezifischen datenschutzrechtlichen Fragen zu adressieren, die sich bei der Verwendung personenbezogener Daten zum Training und Betrieb von KI-Anwendungen stellen. Im Detail besteht jedoch noch Anpassungsbedarf.

I. Art. 9 Abs. 2 lit. k) und Abs. 5 DSGVO-E

Der Kommissionsentwurf COM (2025) 837 final ändert Artikel 9 DSGVO wie folgt:

a) In Absatz 2 werden folgende Buchstaben angefügt:

> *„k) die Verarbeitung erfolgt im Zusammenhang mit der Entwicklung und dem Betrieb eines KI-Systems im Sinne des Artikels 3 Nummer 1 der Verordnung (EU) 2024/1689 oder eines KI-Modells unter den in Absatz 5 genannten Bedingungen,*
> l) […].“

b) Folgender Absatz wird angefügt:

> *„(5) Für die in Absatz 2 Buchstabe k genannte Verarbeitung werden geeignete organisatorische und technische Maßnahmen getroffen, um die Erhebung und sonstige Verarbeitung besonderer Kategorien personenbezogener Daten zu vermeiden. Stellt der Verantwortliche trotz der Umsetzung solcher Maßnahmen fest,*

dass in den für das Trainieren, Testen oder Validieren verwendeten Datensätzen oder im KI-System oder KI-Modell besondere Kategorien personenbezogener Daten enthalten sind, so entfernt er diese Daten. Erfordert das Entfernen dieser Daten einen unverhältnismäßigen Aufwand, so schützt der Verantwortliche diese Daten in jedem Fall unverzüglich wirksam davor, zur Erzeugung von Ergebnissen verwendet, offengelegt oder auf andere Weise Dritten zur Verfügung gestellt zu werden."

1. Die Ergänzung des Art. 9 Abs. 2 DSGVO durch die vorgeschlagene lit. k) ist grundsätzlich zu begrüßen, weil der KI-Anbieter bisher anders als bei (einfachen) personenbezogenen Daten im Rahmen von daran gehindert ist, KI-Systeme mit besonderen Daten i.S.v. Art. 9 DSGVO zu trainieren, sofern nicht ausnahmsweise eine ausdrückliche Einwilligung oder sonstige Erlaubnis nach Art. 9 Abs. 2 DSGVO vorliegt. Die Widerrufsmöglichkeit der Einwilligung ist bei Entwicklung von KI-Systemen und KI-Modellen (v.a. Training) ggf. als teilweise oder vollständige Löschung im Trainingsdatenkorpus umsetzbar. Aber die Löschung im Modell selbst, sofern sich ein Memorization-Effekt auftritt, ist nach aktuellem Stand der Technik kaum praktikabel. Von den sonstigen Erlaubnistatbestände des Art. 9 Abs. 2 DSGVO können insbes. lit. i) und j) im Einzelfall relevant sein – v.a., wenn der Begriff der wissenschaftlichen Forschung erweitert wird (wie in Art. 4 Nr. 38 DSGVO-E vorgesehen). Aber viele Fälle der Entwicklung von KI-Systemen und KI-Modellen durch Unternehmen werden nicht unter diese Erlaubnistatbestände fallen und man kann auch nicht annehmen, dass Entwicklung von KI-Systemen und KI-Modellen per se ein „erhebliches öffentliches Interesse" (Art. 9 Abs. 2 lit. g DSGVO) darstellen. Art. 10 Abs. 5 KI-VO – der durch den „Omnibus für KI" ebenfalls geändert werden soll – ist nur für Hochrisiko-KI-Systeme anwendbar und insoweit auch nur für das Erkennen und Verhindern von Verzerrungen, was als Rechtsgrundlage für die Verarbeitung besonderer Kategorien personenbezogener Daten bei Entwicklung und Betrieb von KI-Systemen und KI-Modellen zu eng ist. Insoweit besteht ein Regelungsbedürfnis.

2. Art. 9 Abs. 5 DSGVO-E setzt voraus, dass geeignete organisatorische und technische Maßnahmen getroffen werden, um die Erhebung und sonstige Verarbeitung besonderer Kategorien personenbezogener Daten zu **vermeiden**. Damit ist Art. 9 Abs. 5 DSGVO-E nur einschlägig für KI-Systeme und Modelle, die für allgemeine Verwendungszwecke bestimmt sind und/oder ihrer Natur nach gerade nicht auf besondere Kategorien personenbezogener Daten angewiesen

sind. Für spezifische Verwendungszwecke, bei denen sowohl für die Entwicklung als auch für den Betrieb besondere Kategorien personenbezogener Daten erforderlich oder unvermeidbar sind, passt Art. 9 Abs. 5 DSGVO-E nicht. Solche KI-Systeme sind auch außerhalb des Hoch-Risiko-Bereichs sehr relevant, nicht nur in der Medizin. Beispiele sind etwa KI-Systeme, die mittels Eye-Tracking-Mechanismus Kindern beim Lesenlernen oder Fremdsprachenunterricht fördern sollen, und bei denen sich aus der Art der Daten Rückschlüsse z.B. Legasthenie ziehen lassen.

3. Um mehr Rechtssicherheit zu erreichen, wären insoweit auch Klarstellungen zumindest in den Erwägungsgründen zu Art. 9 Abs. 2 lit. g) bis j) sowie besser Ergänzungen im Gesetzestext sinnvoll, aus praktischer Sicht wohl sogar notwendig. Klarzustellen wäre insbesondere, dass bereits die Entwicklung, Produktverbesserung und der Betrieb durch Unternehmen der Privatwirtschaft einem öffentlichen oder gar erheblichen öffentlichen Interesse dienen kann.

4. Hinzu kommt, dass KI-Modelle und KI-Systeme für den produktiven Einsatz regelmäßig mit anderer Software, Datenbanken und Hardware verbunden sind (etwa User Frontend, VR-Brillen, Sensoren) und – jedenfalls bei großen Modellen – in einer Rechenzentrums-Infrastruktur betrieben werden. Insoweit stellt sich die Frage, ob die Beschränkung auf das KI-System und das KI-Modell in Art. 9 Abs. 2 lit. k) DSGVO-E für die Praxis tauglich ist, wenn sich die Rechtsgrundlage – jedenfalls könnte man dieses so lesen – nur auf eine technische Komponente eines Gesamtsystems bezieht. Vor diesem Grund empfiehlt sich ein technologieneutralerer Ansatz.

   Das gilt insbesondere auch im Hinblick auf RAG (Retrieval Augmented Generation), also für die Anbindung von sog. Expertenwissen. Dabei werden spezifische Informationen/Referenzdokumente gesammelt (etwa auch durch Crawling spezifischer Webseiten); die Informationen werden vektorisiert und als Vektordatenbank an das KI-System angebunden und ggf. auch in unvektorisierter Form für Quellenangaben gespeichert. Diese Datenbanken sind nicht das KI-System im engeren Sinne. Zwar wird das RAG-System typischerweise als Möglichkeit angesehen, um auch die Rechtmäßigkeit der Datenverarbeitung zu verbessern (DSK, Orientierungshilfe zu datenschutzrechtlichen Besonderheiten generativer KI-Systeme mit RAG-Methode, Version 1.0, Okt. 2025

https://www.datenschutzkonferenz-online.de/media/oh/DSK_OH_RAG.pdf). Allerdings ist auch bei RAG-Systemen – je nach Art und Umfang der Informationen/Referenzdokumente, nicht ausgeschlossen, dass besondere Kategorien personenbezogener Daten – die im Einzelnen für Entwicklung und Betrieb nicht erforderlich sind, mit verhältnismäßigen Mitteln nicht vollständig gefunden und beseitigt werden können.

5. Der bisherige Lösungsansatz in Art. 9 (5) DSGVO-E sieht eine Kaskade an technisch-organisatorischen Maßnahmen vor, die das Ziel verfolgen, die Grundrechte und Grundfreiheiten der betroffenen Person in der Weise zu schützen, als besondere Kategorien personenbezogener Daten

a. im Trainingsdatenkorpus,
b. aber auch in den Validierungs- und Testdaten entfernt sowie
c. im KI-Modell beseitigt werden und,
   soweit das mit verhältnismäßigen Mitteln nicht vollständig möglich ist,
d. im Output oder bei anderen Formen der Offenlegung verhindert werden.

6. Deshalb scheint Art. 9 Abs. 5 DSGVO-E vorrangig auf Entwicklung und Betrieb von solchen KI-Systemen und KI-Modellen (wie LLM) abzuzielen, die mit großen Mengen an Internet-Daten (aus Common Crawl u.ä.) trainiert werden, in denen zwangsläufig und teils für den Anbieter praktisch nicht erkennbar auch Daten im Sinne von Art. 9 Abs. 1 DSGVO enthalten sind, auch wenn diese strikt genommen für die Entwicklung und den Betrieb des KI-Systems oder Modells nicht erforderlich wären. So gesehen ist das stufenweise Vorgehen in Art. 9 Abs. 5 DSGVO-E sinnvoll, wonach die Anbieter/Hersteller angehalten werden, zu prüfen, ob die Erhebung und Verarbeitung besonderer Daten in Trainingsdaten vermieden werden können. Es ist bislang nicht verlässlich möglich, alle Art. 9 DSGVO-Daten in großen Datenmengen automatisiert aufzufinden, da sich z.B. der Charakter als Gesundheitsdatum auch indirekt über den Kontext ergeben kann, wie Art der aufgerufen Webseite (EuGH C-252/21 - Meta Platforms) oder Lieferanschriften, wenn es um bestellte Arzneimittel geht bei (EuGH C-21/23 - Lindenapotheke). Insofern sollte in Art. 9 Abs. 5 DSGVO-E klargestellt werden, dass die Maßnahmenkaskade dem Stand der Technik genügen muss. Es ist auch grundsätzlich richtig, dass insbesondere die Offenlegung von besonderen Kategorien personenbezogener Daten verhindert werden muss. Ob dies jedoch durch ggf. KI-basierte Output-Filter vollständig möglich ist, insbesondere im Hinblick

auf den indirekten Charakter oder Kontext-Bezug, ist jedenfalls nach aktuellem Stand der Technik fraglich. Insoweit wäre zusätzlich für Fälle, in denen Entwicklern, Anbietern oder Betreibern nach dem Stand der Technik keine ausreichend sichere Filterung zur Verfügung steht, ein Melde- und Abhilfeverfahren (ähnlich dem Art. 16 Digital Services Act) denkbar und klarzustellen, dass Sanktionen erst greifen, wenn die entsprechenden technischen Mittel der Filterung nach dem Stand der Technik nicht genutzt wurden. Der Anbieter muss nach Art. 5 Abs. 2 DSGVO Rechenschaft darüber ablegen können, ob eine wirksame Beseitigung von besonderen Kategorien personenbezogener Daten in den Trainings-, Validierungs- und Testdaten sowie im Modell selbst erfolgt ist, insbesondere falls sich im Einzelfall doch solche Daten im Output zeigen sollten. Ggf. müsste also dokumentiert werden, welche Daten bei der Entwicklung oder dem späteren Nachtraining / Finetuning „memoriziert“ bzw. im Output extrahiert werden. Insoweit wäre klarzustellen, inwieweit ggf. zu Dokumentationszwecken Daten, die entfernt wurden, in eingeschränkter Form gespeichert werden müssen.

7. Wichtig wäre klarzustellen, dass die in Art. 9 Abs. 5 DSGVO-E geregelten Maßnahmen insbesondere auch in den Fällen des Art. 9 Abs. 2 lit. g)-j) ein Regelbeispiel für Maßnahmen zur Wahrung der Grundrechte und Interessen der betroffenen Person sein können.

8. Weiter sollte klargestellt werden, dass der Weg über Art. 9 Abs. 5 DSGVO-E nur dann bzw. insoweit beschritten werden kann, wie eine Interessenabwägung faktisch nicht möglich ist, entweder weil die Menge der Daten dies nicht erlaubt (bspw. beim Training von LLMs) oder weil der Kontext der Daten (insbesondere der Charakter von besonderen Daten i.S.v. Art. 9 DSGVO) von Dritten nicht erkannt werden kann; nur dann und soweit das KI-System oder Modell zudem überragenden Interessen der Allgemeinheit dient, wie dies etwa bei LLMs der Fall sein sollte, erscheint es richtig, die Prüfung der Kaskade nach Art. 9 Abs. 5 DSGVO-E an die Stelle einer Interessenabwägung treten zu lassen. Sollte die Abwägung hingegen möglich sein, sollte sie auch bei besonderen Kategorien personenbezogener Daten an die Stelle des Art. 9 Abs. 5 DSGVO-E treten, aber mit dem Verbot eines Outputs der besonderen Daten oder alternativ mit Blick auf den Output entsprechend der für Art. 9 Abs. 5 DSGVO-E angeregten Beschränkungen und Rechte. Die Schwierigkeit dürfe bei diesem Ansatz zugegeben darin liegen, die überragende Bedeutung für die Allgemeinheit ausreichend klar und gleichzeitig nicht so eng zu definieren wie

bisher teils in den Unterfällen des Art. 9 Abs. 2 DSGVO geschehen. Gleichzeitig wird tendenziell aber nur bei einer derartigen Differenzierung ein angemessener Ausgleich zwischen dem Recht auf informationelle Selbstbestimmung einerseits und anderen Freiheitsrechten bzw. Interessen der Allgemeinheit andererseits hergestellt.

9. Diese Kritik zu Art. 9 (5) DSGVO-E gilt entsprechend auch für die vorgeschlagenen Änderungen des Art. 10 EUDPR-E.

10. Die oben ausgeführten Empfehlungen zu einem technologieneutraleren Ansatz gelten nicht nur im KI-Umfeld. In Cloud- und anderen Providermodellen wird Software heute auch sonst mit Hilfe von personenbezogenen Daten entsprechend verbessert. Auch hier kann sich der entsprechende Provider bei besonderen Kategorien personenbezogener Daten regelmäßig auf keine geeignete Rechtsgrundlage berufen, um Produktverbesserung und Produktneuentwicklungen zu legitimieren. Es ist insofern zu überlegen, die Regelungen unter Art. 9 Abs. 2 DSGVO-E in diesem Sinne zu erweitern.

11. Ähnliches gilt für die Fragestellung, wie in Supportfällen mit besonderen Kategorien personenbezogener Daten umgegangen werden darf. Hier agieren entsprechende Software- oder Hardwareanbieter im Grundsatz zwar oftmals als Auftragsverarbeiter, das heißt weisungsgebunden gegenüber den für die Datenverarbeitung Verantwortlichen. Allerdings fehlt es bezogen auf Daten nach Art. 9 Abs. 1 DSGVO gleichwohl oft an einer im konkreten Supportfall erforderlichen Rechtsgrundlage zur Datenverarbeitung. Sollen etwa Fehler einer medizinischen Software behoben werden, mag dieses nur mit Hilfe der Nutzung entsprechender Gesundheitsdaten möglich sein. Dann greifen aber die in Artikel 9 Abs. 2 DSGVO vorgesehenen Ausnahmen im konkreten Fall oftmals nicht. So wird die Situation regelmäßig insbesondere so sein, dass auch die Ausnahme zur Behandlung des Patienten (Art. 9 Abs. 2 lit. c DSGVO) nicht eingreift. Denn nur in seltenen Fällen (so etwa im Rahmen einer Operation) wird die Ad-hoc-Fehlerbeseitigung noch unter diese Ausnahme gefasst werden können. Selbst der Verantwortliche, der die entsprechenden IT-Systeme betreibt, wäre, wenn kein Notfall vorliegt, nicht berechtigt, die Daten zu Fehlerbehebungszwecken zu nutzen. Die Auftragsdatenverarbeitung ist kein Erlaubnistatbestand, sondern stellt den

Auftragsverarbeiter nur so, dass er auf der Basis der Erlaubnis des Verantwortlichen agieren darf.

12. Nicht zuletzt stellt sich als eine der größten Herausforderungen in der Praxis bei Entwicklung und Betrieb von KI-Systemen und KI-Modellen die Erfüllung der Betroffenenrechte dar (v.a. die Informationspflichten nach Art. 14 DSGVO und das Widerspruchsrecht nach Art. 21 DSGVO). Wenn die restriktive Auslegung der Unverhältnismäßigkeit (insbesondere in Art. 14 Abs. 5 lit b DSGVO) nicht korrigiert wird, können die durch Art. 9 Abs. 5 DSGVO-E intendierten Erleichterungen in der Praxis keine Wirkung zeigen. Dazu verweisen wir auf den Vorschlag unten Rn.18.

II. Art. 88c DSGVO-E

Der Kommissionsentwurf COM (2025) 837 final sieht vor:

*„**Artikel 88c Verarbeitung im Zusammenhang mit der Entwicklung und dem Betrieb von KI***

Ist die Verarbeitung personenbezogener Daten im Zusammenhang mit der Entwicklung und dem Betrieb eines KI-Systems im Sinne des Artikels 3 Nummer 1 der Verordnung (EU) 2024/1689 oder eines KI-Modells im Interesse des Verantwortlichen erforderlich, so kann diese Verarbeitung gegebenenfalls aus berechtigtem Interesse im Sinne des Artikels 6 Absatz 1 Buchstabe f der Verordnung (EU) 2016/679 erfolgen, es sei denn, andere Rechtsvorschriften der Union oder der Mitgliedstaaten sehen ausdrücklich eine Einwilligung vor und die Interessen oder Grundrechte und Grundfreiheiten der betroffenen Person, die den Schutz personenbezogener Daten erfordern, überwiegen, insbesondere wenn es sich bei der betroffenen Person um ein Kind handelt.

Eine solche Verarbeitung unterliegt geeigneten organisatorischen, technischen Maßnahmen und Garantien für die Rechte und Freiheiten der betroffenen Person, zum Beispiel um in der Phase der Auswahl der Quellen und des Trainings und Testens von KI-Systemen oder KI-Modellen die Einhaltung der Datenminimierung sicherzustellen, im KI-System oder KI-Modell auf Vorrat gespeicherte Daten vor der Offenlegung zu schützen, für mehr Transparenz für die betroffenen Personen zu sorgen und den betroffenen Personen ein bedingungsloses Recht auf Widerspruch gegen die Verarbeitung ihrer personenbezogenen Daten einzuräumen.“

13. Zwar ist im Grundsatz zu begrüßen, dass klargestellt wird, dass die Nutzung von personenbezogenen Daten für die Entwicklung und den Betrieb von KI-Systemen und KI-Modellen im Rahmen des Artikel 6 Abs. 1 S. 1 lit. f DSGVO erlaubt ist. Dafür bedarf es allerdings keiner neuen Norm. Vielmehr wäre z.B. eine Ergänzung entsprechender Beispiele in den Erwägungsgründen zu Artikel 6 Abs. 1 S. 1 lit. f DSGVO ausreichend. Anderenfalls zerfasern die Regelungen der DSGVO, deren Text dann unnötig lang wird. Es besteht die Gefahr, dass in unbedeutende Abweichungen der Regelungen Unterschiede hineingelesen werden, die gar nicht intendiert waren. Zudem entstehen mitunter zwei Regelungsregime der Prüfung in der Rechtspraxis, die zusätzliche Rechtsunsicherheit und unnötige Kosten erzeugt, insbesondere wenn aufgrund der technologiespezifischen Regelung in Art. 88c DSGVO-E – wie oben unter 3.-4. und 10.-11. dargestellt – für ein Gesamtsystem bzw. für einen einheitlichen Lebenssachverhalt und Zweck verschiedene Rechtsgrundlagen angewendet werden müssen.

14. Es stellt sich bei einer solchen Spezialregelung die Frage, ob nicht für jede neue Technologie oder jedes neue Produkt eine gesonderte Regelung für die Abwägung im Rahmen von Artikel 6 Abs. 1 S. 1 lit. f DSGVO eingeführt werden müsste. Dieses ist allerdings nicht erforderlich, weil die Normen der DSGVO im Grundsatz technologieoffen gestaltet sind.

15. Das Gleiche gilt auch für den zweiten Absatz des Entwurfs des Artikel 88 c DSGVO. Auch dieser entspricht weitestgehend der heutigen Gesetzeslage. Will man dieses betonen, wird auch hier vorgeschlagen, die entsprechenden Aspekte in einem Erwägungsgrund zu betonen.

16. Was jedoch bei Entwicklung von LLM und anderen Big Data-Technologien bislang nicht ausreichend in Art. 6 Abs. 1 lit. f) DSGVO adressiert ist, ist der Umstand, den Art. 9 Abs. 5 DSGVO-E in den Fokus nimmt (vgl. Rn. 6.), dass nämlich ggf. zwingend mit großen Datenmengen aus dem Internet entwickelt (z.B. trainiert) werden muss und eine detaillierte Interessenabwägung ausgerichtet an einzelnen Datenarten und spezifischen Interessen von betroffenen Personen faktisch nicht möglich ist – jedenfalls nicht in einem Detailgrad, wie es die sehr restriktive Rechtsprechung des EuGH implizit verlangt (etwa EuGH Urt. v. 4.10.24 - C-621/22; EuGH Urt. v. 12.09.2024 - C-17/22, C-18/22) und wie es von den

Datenschutzbehörden gefordert wird (siehe EDSA, Leitlinien 1/2024[1]). Insoweit wäre eine Klarstellung im Anwendungsbereich des Art. 6 Abs. 1 lit. f) DSGVO dahingehend wichtig, dass wirksame technische und organisatorischen Maßnahmen (vergleichbar mit oder entsprechend Art. 9 Abs. 5 DSGVO-E) einen verhältnismäßigen Interessenausgleich schaffen können, soweit das KI-System oder Modell überragenden Interessen der Allgemeinheit dient (dazu Rn. 8).

17. Regelmäßig werden KI-Systeme und KI-Modelle durch den Anbieter trainiert, und zwar auf Basis von Altdatenbeständen und/oder von durch Kunden zur Verfügung gestellten Daten. Geschieht dieses, verarbeiten entsprechende Anbieter die Daten nicht mehr zu Zwecken des Kunden, sondern (zumindest auch) zu eigenen Zwecken und sind insoweit nicht mehr rein weisungsabhängige Auftragsverarbeiter im Sinn von Artikel 28 DSGVO. Die Produktverbesserung im Rahmen der Auftragsverarbeitung ist grundsätzlich keine Datenverarbeitung zu eigenen Zwecken des Kunden, wie sowohl die deutsche DSK im Kontext der Nutzung von Telemetriedaten für Microsoft 365 [„Bewertung der aktuellen Vereinbarung zur Auftragsverarbeitung“][2] entschieden haben als auch die französische Datenschutzbehörde CNIL [„Determining the legal qualification of AI system providers“][3]. Anders kann es sein, wenn die Produktverbesserung nur für und im Auftrag individueller Kunden erfolgt und jeweils nur mit deren personenbezogenen Daten, und wenn die Verbesserung/Weiterentwicklung nicht in den „Standard“ des Anbieters übernommen wird. Soweit keine Auftragsverarbeitung vorliegt, müssen sich die Anbieter als Verantwortliche insofern selbst auf Artikel 6 bzw. Artikel 9 DSGVO (ggf. in seiner dann künftig erweiterten Fassung) unter Berücksichtigung der Zweckänderung nach Art. 6 Abs. 4 DSGVO stützen können. Das hat die weiteren nachstehend beschriebenen, die Nutzung von KI tendenziell behindernden Folgen.

18. Verarbeiten Auftraggeber und Auftragnehmer die Daten, die zunächst Gegenstand der Auftragsverarbeitung sind, mit gleichen Mitteln, nämlich dem KI-Angebot des Auftragnehmers weiter, kann in Frage stehen, ob sich die Verantwortung der Parteien noch abgrenzen lässt oder ob eine gemeinsame Verantwortung i.S.v. Art. 26 DSGVO vorliegt und ob diese v.a. bei Behörden als Kunden überhaupt zulässig

[1] EDPB Guidelines 1/2024 on processing of personal data based on Article 6(1)(f) GDPR, Version 1, adopted on 8 October 2024, abrufbar hier: edpb_guidelines_202401_legitimateinterest_en.pdf
[2] Vgl. https://www.datenschutzkonferenz-online.de/media/dskb/2022_24_11_festlegung_MS365_abschlussbericht.pdf.
[3] Vgl. unter https://www.cnil.fr/fr/node/164396.

gestaltet werden kann (siehe Art. 6 Abs. 1 UAbs. 2 DSGVO). Diese gemeinsame Verantwortung wird von Kunden wie von Anbietern aber regelmäßig u.a. wegen der Rechtsfolgen und Haftungsrisiken abgelehnt. Es ist insofern zu erwägen, klarzustellen (was Art. 4 Nr. 7 DSGVO als Möglichkeit auch vorsieht), dass Kunde und Anbieter nicht zu gemeinsam Verantwortlichen nach Art. 26 DSGVO werden, sondern jeweils allein Verantwortliche sind, wenn bei hinreichender Transparenz und Wahlmöglichkeit für den Kunden mit dessen Einverständnis z.B. RAG-Systeme (insbesondere Wissens- und Vektordatenbanken) sowie Finetuning zur Verbesserung von KI-Systemen durch den Anbieter zur Verbesserung und Weiterentwicklung von seinem KI-Angebot verwendet werden. Ggf. kann das in den Erwägungsgründen erfolgen.

19. Aus praktischer Sicht problematisch ist überdies, dass die betroffenen Personen über diese Art der Datenverarbeitung bei Einwicklung und Produktverbesserung über die Zweckänderung und Verarbeitung zu eigenen Zwecken eines (neuen) zusätzlichen Verantwortlichen informiert werden müssen, was de facto oft kaum möglich ist. Insbesondere sehen die Änderungen des „Digital Omnibus zur DSGVO“ zwar Erleichterung bei Art. 13 DSGVO-E, aber nicht bei Art. 14 DSGVO vor. Hier sollte es eine Regelung geben, die entsprechende Informationen im Rahmen von Artikel 13 Abs. 3 und 14 Abs. 4 nur im Rahmen des Verhältnismäßigen fordert und klarstellt, dass diese Regelung entgegen der bisherigen Praxis nicht zu eng ausgelegt wird. Soweit betroffene Personen nicht identifiziert werden können, gelten Art. 11 Abs. 2 und Art. 12 Abs. 2 DSGVO. Aber auch wenn eine Identifizierung möglich wäre, scheitert eine Information nach Art. 14 DSGVO häufig an einer elektronischen Direktkontaktmöglichkeit (per E-Mail). Erwägungsgrund 58 Satz 2 der DSGVO geht hinsichtlich der Informationspflichten (Art. 12-14 DSGVO) davon aus, dass diese Informationen „*in elektronischer Form bereitgestellt werden [können], beispielsweise auf einer Website, wenn sie für die Öffentlichkeit bestimmt ist.“* Allerdings müssen die Informationen für die betroffene Person leicht zugänglich sein. Wenn die Daten nicht direkt bei der betroffenen Person erhoben werden, kann von dieser nicht verlangt werden, dass sie regelmäßig ohne Anhaltspunkte das Internet nach Art. 14 DSGVO-Informationen durchsucht. Insoweit wird sich regelmäßig die Frage stellen, ob und wie betroffene Personen davon erfahren, dass Daten über sie, die im Internet öffentlich zugänglich sind, zum Trainieren, Validieren, Testen von KI oder für RAG-Systeme verwendet werden. Zu denken wäre an eine Begrenzung der Informationspflicht auf die elektronische Bereitstellung

(beispielsweise auf Webseiten), soweit für die Anbieter / Betreiber der KI-Systeme und KI-Modelle keine elektronische Direktkommunikation mit den betroffenen Personen möglich ist. Ist die Direktkommunikation dagegen elektronisch möglich, sollte in Art. 14 Abs. 3 DSGVO klargestellt werden, dass die Anforderungen nach lit. a, b und c alternativ und nicht kumulativ gelten. D.h. wenn die Kommunikation mit der betroffenen Person konkret und verbindlich vorgesehen ist (Art. 14 Abs. 3 lit. b DSGVO), kann die Höchstfrist des Art. 14 Abs. 3 lit. a DSGVO (ein Monat) überschritten werden.

Eine entsprechende Ergänzung sollte bei Art. 14 Abs. 5 lit. b DSGVO erfolgen.

Teil 2: Änderungsvorschläge in Bezug auf die KI-Verordnung

Der DAV begrüßt die auch die vorgesehene Änderung in Artikel 4a KI-VO, möchte jedoch die Gelegenheit nutzen, um auf folgende Problempunkte aufmerksam zu machen:

I. Artikel 4a KI-VO-E

Der Kommissionsvorschlag COM(2025) 836 final sieht vor:

*"**Article 4a Processing of special categories of personal data for bias detection and mitigation***

1. *To the extent necessary to ensure bias detection and correction in relation to high-risk AI systems in accordance with Article 10 (2), points (f) and (g), of this Regulation, providers of such systems may exceptionally process special categories of personal data, subject to appropriate safeguards for the fundamental rights and freedoms of natural persons. In addition to the safeguards set out in Regulations (EU) 2016/679 and (EU) 2018/1725 and Directive (EU) 2016/680, as applicable, all the following conditions shall be met in order for such processing to occur:*
   (a) the bias detection and correction cannot be effectively fulfilled by processing other data, including synthetic or anonymised data;
   (b) the special categories of personal data are subject to technical limitations on the re-use of the personal data, and state-of-the-art security and privacy-preserving measures, including pseudonymisation;

(c) the special categories of personal data are subject to measures to ensure that the personal data processed are secured, protected, subject to suitable safeguards, including strict controls and documentation of the access, to avoid misuse and ensure that only authorised persons have access to those personal data with appropriate confidentiality obligations;

(d) the special categories of personal data are not transmitted, transferred or otherwise accessed by other parties;

(e) the special categories of personal data are deleted once the bias has been corrected or the personal data has reached the end of its retention period, whichever comes first;

*(f) the records of processing activities pursuant to Regulations (EU) 2016/679 and (EU) 2018/1725 and Directive (EU) 2016/680 include the reasons why the processing of special categories of personal data **was necessary** to detect and correct biases, and why that objective could not be achieved by processing other data.*

2. Paragraph 1 may apply to providers and deployers of other AI systems and models and deployers of high-risk AI systems where necessary and proportionate if the processing occurs for the purposes set out therein and provided that the conditions set out under the safeguards set out in this paragraph.

20. Im Hinblick auf Art. 4a KI-VO-E ist zu begrüßen, dass das Verhältnis zwischen einerseits hochwertige Trainingsdaten anstrebenden Data Governance und der Vermeidung von Verzerrungen, mithin Bias-Problemen (bisher Art. 10 KI-VO) und andererseits dem DSGVO-Grundsatz der Datenminimierung geschärft und im Detail in Abs. 1 S. 1 und in lit. f neu feinjustiert werden soll. Ebenso ist zu begrüßen, dass dies nicht bloß für Hochrisiko-Systeme (Art. 10 Abs. 1 KI-VO-Entwurf) geschieht und daher Abs. 2 neu eingeführt wird, der Art. 4a KI-VO-Entwurf unter bestimmten Umständen zusätzlich für (normale) KI-Systeme und KI-Modell und die Betreiber von Hochrisiko-KI-Systemen zur Anwendung bringt.

21. Soweit insofern in Art. 4a Abs. 1 S. 1 KI-VO-E für die Entdeckung und Korrektur von Verzerrungen (Bias) darauf verzichtet wird, dass dieses strikt erforderlich („*strictly necessary*“) sind, und stattdessen nur auf eine (einfache) Erforderlichkeit abgestellt wird, ist diese Gesetzesänderung für Hochrisiko-Systeme jedenfalls zu begrüßen. Wie schon im Rahmen von Art. 10 Abs. 5 KI-VO (in der bisherigen Fassung) erscheint die in Abs. 1 S. 1 des Art. 4a KI-VO-E gewählte Formulierung „*subject to*

appropriate safeguards for the fundamental rights and freedoms of natural persons“ zu unbestimmt. Auch wenn damit ein Auffangtatbestand geschaffen werden soll, der durch S. 2 nur exemplarisch ergänzt wird, wäre es für den Rechtsanwender hier wünschenswert, weitere Leitlinien zur Auslegung zu erhalten.

22. Mit Blick auf den Katalog des Art. 4a Abs. 1 lit. a) bis f) KI-VO-E, der teilweise allgemeine DSGVO-Grundsätze aus Art. 5, 25, 32 DSGVO regelt, ist nicht klar, inwieweit neben der Erfüllung des Katalogs noch weitere „safeguards“ hinzukommen müssen. Der Verweis auf die DSGVO müsste konkretisiert werden, auch vor dem Hintergrund, dass die DSGVO parallel modifiziert wird, wie das im Grundsatz durch Verordnungsvorschlag COM (2025) 837 final vorgesehen ist (dazu siehe oben).

23. Auch sind die weiter aufgeführten zusätzlichen Bedingungen, wie schon im Rahmen der aktuellen Fassung von Art. 10 Abs. 5 KI-VO, zu hinterfragen:

    (a) Lit. a des Entwurfs regelt wie in der aktuellen Fassung von Art. 10 Abs. 5 KI-VO, eigentlich eine Selbstverständlichkeit, die sich schon aus dem Grundsatz der Erforderlichkeit ergibt („*To the extent necessary to ensure bias detection and correction …*“) sowie aus dem Grundsatz der Datenminimierung nach Art. 5 Abs. 1 lit. c DSGVO. Dass zuvörderst synthetische oder anonymisierte Daten genutzt werden müssen, ist nach der Behördenpraxis zur Entwicklung und zum Testen von IT-Systemen insofern ein anerkannter Grundsatz.

    (b) Lit. b des Entwurfs ist, wie schon in der aktuellen Fassung von Art. 10 Abs. 5 KI-VO, nach wie vor unklar abgefasst. Insbesondere wird nicht deutlich, ob die Regelung eine Verschärfung oder Erleichterung gegenüber den Vorgaben des Art. 5 Abs. 1 lit. b und des Art. 32 DSGVO darstellen soll und, wenn ja, worin diese liegt.

    (c) Auch lit. c des Entwurfs wiederholt, wie schon in der aktuellen Fassung von Art. 10 Abs. 5 KI-VO, im Wesentlichen die aus Art. 32 DSGVO ohnehin bekannten Grundsätze. Soweit durch die Regelbeispiele der „*controls*“ und Dokumentation zusätzliche Maßnahmen gefordert werden, wäre es hilfreich, dies in Abgrenzung zu Art. 32 DSGVO zu betonen.

(d) Die redaktionelle Änderung in lit. d) des Entwurfs im Vergleich zur aktuellen Fassung, nämlich die Streichung von „*to be*“, wird begrüßt. Allerdings fragt sich auch bei lit. d), ob diese Bestimmung mit Blick auf die Grundsätze der Erforderlichkeit und Datenminimierung („*To the extent necessary to ensure bias detection and correction …“*) entbehrlich ist.

(e) Bei dem (der aktuellen Fassung wortgleichen) Entwurf der lit. e) ist noch einmal kritisch zu hinterfragen, ob nicht die Trainingsdatenbasis zur Korrektur des Bias in einigen Fällen erhalten bleiben sollte, um später bei einem Fehlverhalten des KI-Systems, den Grund hierfür prüfen zu können. Das könnte insbesondere für medizinische KI-System gelten.

(f) Lit. f) des Entwurfs reflektiert mit der Streichung des „*strictly*“ die zu begrüßende Absenkung der Anforderungen des Abs. 1 S. 1 des Entwurfs. Auch bei lit. f) lässt sich fragen, ob die Dokumentationspflicht sich nicht schon aus Art. 5 Abs. 2 DSGVO ergibt. Da diese aber zweifelhaft ist und die Dokumentation auch zur Reflexion der Rechtsfragen sinnvoll erscheint, begrüßen wir die ausdrückliche Betonung.

24. Die Regelung in Art. 4a Abs. 2 KI-VO-E ist unklar. So fragt sich, ob das „*where necessary and proportionate*” sich auf die Notwendigkeit der Bias-Detektion oder -Korrektur beziehen soll oder auf die zu ergreifenden Maßnahmen und Sicherungen, die ggf. zugunsten der Anbieter und Entwickler reduziert werden sollten. In Ansehung der Tatsache, dass es dann weiter heißt „...*provided that the conditions set out under the safeguards set out in this paragraph.*”, gehen wir davon aus, dass eigentlich zum Ausdruck gebracht werden soll, dass besondere Kategorien personenbezogener Daten nur dann zur Erkennung und Korrektur von Verzerrungen (Bias) bei einfachen (nicht-Hochrisiko-) KI-Systemen und KI-Modellen sowie von Betreibern von Hochrisiko-KI-Systemen genutzt werden dürfen, wenn dies notwendig und verhältnismäßig ist. Hier ist nachvollziehbar, dass wegen des tendenziell geringeren Gefahrenpotenzials von (normalen) KI-Systemen eine Abwägung im Hinblick auf die Verhältnismäßigkeit der Datenverarbeitung und der geforderten Maßnahmen risikobasiert vorzunehmen ist. Fraglich ist aber, wie dies bei KI-Modellen mit allgemeinem Verwendungszweck erfolgen soll, bei denen der spätere Einsatzzweck regelmäßig für die Anbieter nicht feststeht oder nicht bekannt ist. Überdies stellt sich das oben unter 16 diskutierte Problem, dass gerade bei

LLMs und KI-Systemen mit Massendaten-Training eine Abwägung im Sinne von Art. 6 Abs. 1 S. 1 lit. f DSGVO nicht möglich ist. Insgesamt ist die Regelung klarer zu fassen.

25. Es fragt sich weiter mit Blick auf Art. 4 a Abs. 1 und 2 KI-VO-E, warum diese auf Daten nach Art. 9 Abs. 1 DSGVO beschränkt sind und im Übrigen nicht andere personenbezogenen Daten adressieren. Bei diesen ist die Interessenabwägung mitunter nicht durchführbar (siehe oben 16.).

26. Die Änderung in Art. 10 Abs. 1 KI-VO ist folgerichtig.

Verteiler

Europa

Europäische Kommission
- Generaldirektion Kommunikationsnetze, Inhalte und Technologien (DG CNECT)
- Generaldirektion Justiz und Verbraucher (DG JUST)

Europäisches Parlament
- Rechtsausschuss (JURI)
- Ausschuss für Binnenmarkt und Verbraucherschutz (IMCO)
- Ausschuss für Bürgerliche Freiheiten, Justiz und Inneres (LIBE)

Rat der Europäischen Union
Ständige Vertretung der Bundesrepublik Deutschland bei der EU
Justizreferenten der Landesvertretungen bei der EU
Rat der Europäischen Anwaltschaften (CCBE)
Bundesverband der Freien Berufe (BFB) – Büro Brüssel
Deutsche Industrie- und Handelskammer (DIHK) – Büro Brüssel
Bundesverband der deutschen Industrie e.V. (BDI) – Büro Brüssel

Deutschland

Bundesministerium des Innern
Bundesministerium der Justiz und für Verbraucherschutz
Bundesministerium für Wirtschaft und Energie
Bundesministerium für Digitales und Staatsmodernisierung

Ausschuss für Inneres im Deutschen Bundestag
Ausschuss für Recht und Verbraucherschutz im Deutschen Bundestag
Ausschuss für Wirtschaft und Energie im Deutschen Bundestag
Ausschuss Digitale Agenda im Deutschen Bundestag
Fraktionen im Deutschen Bundestag

Die Bundesbeauftragte für den Datenschutz und die Informationsfreiheit

Die Justizministerien der Länder
Die Datenschutzbeauftragten der Bundesländer

Europäische Kommission - Vertretung in Deutschland
Bundesrechtsanwaltskammer
Bundesnotarkammer
Bundesverband der Freien Berufe e.V.
Deutscher Richterbund, Bund der Richterinnen und Richter, Staatsanwältinnen und
Bund Deutscher Verwaltungsrichter und Verwaltungsrichterinnen
Staatsanwälte e.V. (DRB)
Deutscher Notarverein
Deutscher Steuerberaterverband e.V. Berlin
Bundesverband der Deutschen Industrie e.V.
Arbeitsgemeinschaft berufsständischer Versorgungseinrichtungen e.V.
Deutscher EDV-Gerichtstag e.V.
GRUR - Deutsche Vereinigung für gewerblichen Rechtsschutz und Urheberrecht e.V.
Bitkom e. V.
Deutsche Gesellschaft für Recht und Informatik e.V. (DGRI)
ver.di - Vereinte Dienstleistungsgewerkschaft

Gewerkschaft der Polizei
Deutsche Polizeigewerkschaft im DBB (DPolG)

DAV-Vorstand und Geschäftsführung
Vorsitzende der DAV-Gesetzgebungsausschüsse
Vorsitzende der DAV-Landesverbände
Vorsitzende des FORUMs Junge Anwaltschaft

Presse

Frankfurter Allgemeine Zeitung GmbH
Süddeutsche Zeitung GmbH
Redaktion NJW
JUVE Verlag für juristische Information GmbH
Redaktion Legal Tribune Online / LTO
Redaktion Anwaltsblatt
juris GmbH
Redaktion MultiMedia und Recht (MMR)
Redaktion Zeitschrift für Datenschutz ZD
Redaktion heise online
DER SPIEGEL GmbH & Co. KG
Computer und Recht