Stellungnahme

der

TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e. V.

zu den

Guidelines 01/2025 on Pseudonymisation des European Data Protection Board vom 16. Januar 2025

Berlin, 14.03.2025

Zusammenfassung

Pseudonymisierung ist eine in der Forschung weit verbreitete und akzeptierte Maßnahme um Risiken der Verarbeitung personenbezogener Daten zu reduzieren, ohne die erforderlichen Anwendungsfälle bzw. den Nutzen der Verarbeitung einzuschränken. Aufgrund der besonderen Sensibilität der in der medizinischen Forschung benötigten Gesundheitsdaten kommt dieser Maßnahme hier eine besondere Relevanz zu. In der DSGVO wird Pseudonymisierung zudem definiert und hat damit als mögliche Schutzmaßnahme bei der Verarbeitung personenbezogener Daten einen herausgehobenen Stellenwert. Insofern ist grundsätzlich zu begrüßen, dass der Europäische Datenschutzausschuss (EDSA) eine Leitlinie dazu veröffentlicht, was unter einer Pseudonymisierung im Sinne der DSGVO zu verstehen ist und welche Grenzen bzw. Abgrenzungen es diesbezüglich gibt.[1]

Pseudonymisierung ist in der DSGVO in Art. 4 Nr. 5 einerseits als eine bestimmte Art der Verarbeitung personenbezogener Daten und andererseits als eine Verarbeitung mit einer bestimmten Zielsetzung definiert.

Bezüglich der **Art der Verarbeitung bei einer Pseudonymisierung** beschreibt die DSGVO in der Begriffsbestimmung in Art. 4 Nr. 5 sehr klar, dass es hier um die Trennung von Informationen geht und ein Teil der Informationen, die offensichtlich einer Identifizierung betroffener Personen dienen können, bestimmten technischen und organisatorischen Maßnahmen unterliegen. Die Leitlinie des EDSA analysiert zwar die Vorgaben zur Art der Verarbeitung im Rahmen einer Pseudonymisierung, verkennt aber, dass es hierbei um einen Wesenskern der Pseudonymisierung und somit notwendige Eigenschaften einer Pseudonymisierung geht. Insofern werden an vielen Stellen in der Leitlinie Maßnahmen zur Umsetzung einer Pseudonymisierung vorgeschlagen, die hinsichtlich der Art der Verarbeitung die notwendigen Bestandteile einer Pseudonymisierung gar nicht aufweisen. Einige der vorgeschlagenen Maßnahmen gehören vom Charakter her eher in den Bereich der Löschung bzw. wären im Rahmen einer Anonymisierung anwendbar, haben aber nichts mit der Abtrennung bestimmter Informationen und deren Weiterverarbeitung in geschützter Form zu tun.

Die **Zielsetzung einer Pseudonymisierung** ist in Art. 4 Nr. 5 DSGVO so beschrieben, dass man nur bei Berücksichtigung dieser Definition davon ausgehen muss, dass der Teil der Daten, der nach der Abtrennung der Daten mit besonderem Identifikationspotential übrigbleibt, so gut wie anonym sein müsste. Eine ausführliche gutachterliche Stellungnahme, die die TMF hierzu vor einigen Jahren eingeholt hat, kommt zu dem Ergebnis, dass für ein umfassendes Verständnis des Konzepts der Pseudonymisierung in der DSGVO die Heranziehung der Definition in Art. 4 Nr. 5 alleine nicht ausreichend ist, sondern vielmehr alle Verwendungen bzw. Verweise auf das Konzept in der DSGVO berücksichtigt werden müssen [1]. Demnach kann das Ergebnis einer Pseudonymisierung durchaus sein, dass ein Teil der Daten nach einer Abtrennung bestimmter Informationen keinen Personenbezug mehr aufweist. In aller Regel ist aber das Ergebnis ein weiterhin personenbezogener Teildatensatz, der lediglich im Sinne einer Datenminimierung so reduziert ist, dass sich in diesem Teildatensatz keine Daten mehr befinden, die für die angestrebten Ziele der Verarbeitung nicht erforderlich sind. Anstatt zwei

[1] siehe https://www.edpb.europa.eu/our-work-tools/documents/public-consultations/2025/guidelines-012025-pseudonymisation_en

mögliche und sich deutlich unterscheidende Zielsetzungen der Pseudonymisierung zu unterscheiden, folgt die Leitlinie des EDSA eher einem unscharfen Mittelweg, in dem sie suggeriert, dass man mit einer Pseudonymisierung den nach der Abtrennung der offensichtlich identifizierenden Daten übriggebliebenen Teil möglichst nahe an einen anonymen Datensatz heranführen müsste, ohne aber Anonymität notwendigerweise zu erreichen. Die Leitlinie des EDSA interpretiert insofern die Zielsetzung einer Pseudonymisierung viel zu weitgehend und schlägt dann zur Erreichung dieser Zielsetzung Maßnahmen vor, die von der Art der Verarbeitung her von der Definition der Pseudonymisierung nicht mehr gedeckt sind. Diese Vorgabe ist aber weder für die Anwender einer Pseudonymisierung hilfreich, die damit im Sinne der Datenminimierung handeln wollen, ohne aber Anwendungsfälle oder die Datenqualität einzuschränken, noch für die Anwender, die im Ergebnis einer Pseudonymisierung möglicherweise tatsächlich im rechtlichen Sinne nicht mehr personenbeziehbare Daten ansteuern. Hier droht die Verschiebung der Zielsetzung der Pseudonymisierung in Richtung einer Anonymisierung die Maßnahme gerade für die Forschung, in der diese Maßnahme heute noch weit verbreitet und breit akzeptiert ist, unattraktiv zu machen. Nicht auszuschließen, dass in Folge einer solchen Umdeutung selbst die basale und in der Praxis sehr hilfreiche Maßnahme der Abtrennung direkt identifizierender Daten seltener als bisher Anwendung findet. Gesetzliche Vorgaben zur Pseudonymisierung auf nationaler oder auch europäischer Ebene könnten zudem viele Anwendungsfälle der wissenschaftlichen Forschung künftig ausschließen.

Da die Pseudonymisierung gemäß Definition in der DSGVO ein Vorgang der Aufteilung von Daten mit unterschiedlichen Eigenschaften hinsichtlich der Identifizierbarkeit und deren getrennter Weiterverarbeitung ist, stellt sich hier in besonderer Weise die schon seit langem kontrovers diskutierte Frage danach, ob man das **Konzept des Personenbezugs**, welches der DSGVO zugrunde liegt, als ein relatives oder absolutes verstehen muss. Da der EuGH zu dieser Frage schon Stellung bezogen hat und zu dem Ergebnis gekommen ist, dass der DSGVO ein relatives Verständnis zugrunde liegt [2; 3], ist es bedauerlich, dass sich die Leitlinie des EDSA hierzu nicht eindeutig positioniert. Vielmehr gibt es in der Leitlinie viele Formulierungen, die ein absolutes Verständnis nahelegen und damit den Möglichkeitsraum der Pseudonymisierung hinsichtlich unterschiedlicher Zielstellungen unnötig einengen.

Darüber hinaus enthält die Leitlinie viele rechtlich und technisch korrekte Feststellungen, die für unterschiedliche Anwender durchaus hilfreich sein können. Da sie aber in ein unscharfes Verständnis der Pseudonymisierung sowohl hinsichtlich der Art als auch der Ziele der Verarbeitung eingebettet sind, wird der Nutzen der Leitlinie stark eingeschränkt. Gerade in der Forschung würde die Bedeutung der Pseudonymisierung drastisch abnehmen, sollte sich das in der Leitlinie dargelegte Verständnis durchsetzen, da viele der beschriebenen Maßnahmen erfahrungsgemäß mit einem deutlichen Qualitätsverlust der Daten einhergehen würden. Insofern ist eine grundlegende Überarbeitung der Leitlinie im Sinne der hier vorgelegten Analyse dringend erforderlich.

Die Art der Verarbeitung personenbezogener Daten im Rahmen einer Pseudonymisierung

Die Pseudonymisierung ist in der DSGVO in Art. 4 Nr. 5 definiert:

> *'pseudonymisation' means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person;*

In dieser Definition sind Hinweise auf eine bestimmte Art der Verarbeitung und die Zielsetzung der Verarbeitung eng miteinander verwoben. Wenn man nur den Teil der Definition betrachtet, der auf die Art der Verarbeitung eingeht, dann ist hier zum einen von separat gehaltenen Informationen und zum anderen von der Anwendung technischer und organisatorischer Schutzmaßnahmen, denen diese separat gehaltenen Informationen unterliegen, die Rede. Diese Hinweise auf eine bestimmte Art der Verarbeitung sind in ihrer hier gewählten Knappheit offensichtlich unvollständig. Damit bestimmte Informationen separat gehalten werden können, muss zuerst eine Aufteilung von Informationen und sodann eine Abtrennung stattfinden.

Die Zuordnung der Daten zu den betroffenen Personen soll nach der Abtrennung bestimmter Informationen ohne die Nutzung dieser abgetrennten Informationen nicht mehr möglich sein. Das lässt allerdings offen, dass eine Zuordnung mit Nutzung der abgetrennten Informationen durchaus möglich sein kann. In diesen Fällen muss die Verarbeitung so erfolgen, dass eine Verbindung zwischen den abgetrennten Informationen und den eigentlich zu nutzenden Daten bestehen bleiben bzw. etabliert werden muss. Diese Verbindung entsteht typischerweise durch die Nutzung von Pseudonymen, die in der Definition ebenfalls keine Erwähnung finden. Es findet dann also eine Ersetzung von abgetrennten Informationen durch ein Pseudonym statt.

Und die Definition lässt weiter offen, welcher Art die technischen und organisatorischen Schutzmaßnahmen sein sollen, die auf die abgetrennten Informationen angewendet werden. In der Praxis findet man hier beispielsweise einerseits den Anwendungsfall, dass die abgetrennten Informationen geschützt in einer Vertrauensstelle gehalten und verarbeitet werden sowie andererseits die Anwendung einer kryptographischen Transformation auf die abgetrennten Informationen. Diese kryptographische Transformation kann als reversible Verschlüsselung oder auch als Einweg-Verschlüsselung ausgestaltet sein. In allen Fällen ist rein technisch eine Reversibilität der Pseudonymisierung möglich. Im Fall einer Einweg-Verschlüsselung befindet sich der Schlüssel hierzu zu einem Teil beim pseudonymisierenden Verantwortlichen oder Auftragsverarbeiter, der das Verfahren und die eingesetzten Parameter der Verschlüsselung kennt, und zum anderen Teil bei der betroffenen Person, die die in die Verschlüsselung eingehenden Daten kennt.

Die Leitlinie des EDSA enthält im Abschnitt 2 „Definitions and legal analysis" nur eine unvollständige und unscharfe Analyse der Art der Verarbeitung personenbezogener Daten im Rahmen einer Pseudonymisierung (vergl. Abs. 18–20). Die Generierung und Nutzung zusätzlicher Informationen soll demnach ein inhärenter Teil der Pseudonymisierung sein (Abs. 19), was offenließe, dass in Ergänzung auch ganz andere Verarbeitungen zu einer Pseudonymisierung gehören könnten. Diese unscharfe Darstellung verkennt, dass die Abtrennung von Informationen und die Anwendung bestimmter Schutzmaßnahmen auf diese Informationen ein notwendiges Wesensmerkmal der Pseudonymisierung darstellen. Eine

Verarbeitung von Daten, die sich in keiner Weise auf eine solche Abtrennung von Informationen und die Anwendung von Schutzmaßnahmen auf diese abgetrennten Informationen bezieht, ist demnach keine Pseudonymisierung. Insofern ist die Beschreibung der Pseudonymisierung in der Leitlinie unscharf und infolgedessen können die hier beschriebenen Eigenschaften einer Pseudonymisierung nicht zur Abgrenzung des Begriffs gegenüber anderen Verarbeitungen genutzt werden.

In Abschnitt 3.1 enthält die Leitlinie dann eine ausführlichere Beschreibung pseudonymer Transformationen, die ausführlich auf die Abtrennung, Ersetzung und den Schutz der abgetrennten Daten eingeht. Auch hier wird nicht gesehen, dass die Pseudonymisierung nur bestimmte Arten der Verarbeitung umfasst. So findet sich in Abs. 84 die Aussage:

> *[…] Insofar as necessary for pseudonymisation to have the intended effect, it also modifies other attributes, e.g. by removal, generalisation and noise addition.*

Hier werden also als Teil einer Pseudonymisierung Maßnahmen wie die Löschung, die Generalisierung oder auch das Verrauschen von Informationen vorgeschlagen, die in keiner Weise mehr einen Bezug zu den Maßnahmen der Abtrennung und Ersetzung bestimmter identifizierender Informationen sowie deren Schutz haben. Solchen über die Pseudonymisierung hinausschießenden Maßnahmen, widmet die Leitlinie dann noch den ganzen Abschnitt 3.1.3 „Modification of original data necessary for the objectives of pseudonymisation". Es wird dabei deutlich, dass einerseits in der Leitlinie übersehen wird, dass die Pseudonymisierung auch von der Art der Verarbeitung her definiert ist und insofern dazu unpassende Verarbeitungen nicht Teil einer Pseudonymisierung sein können. Andererseits wird dieser überschießende Teil als notwendig zur Erreichung des gesetzlich vorgeschriebenen Ziels einer Pseudonymisierung dargestellt, so dass auch diese Zielsetzung einer genauen Analyse bedarf.

Das Ziel einer Pseudonymisierung

Die Zielsetzung einer Pseudonymisierung ist in Art. 4 Nr. 5 DSGVO so beschrieben:

> *[…] processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information […]*

Zudem soll die „additional information" mit Hilfe technischer und organisatorischer Verfahren so geschützt werden, dass sichergestellt ist

> *[…] that the personal data are not attributed to an identified or identifiable natural person;*

Berücksichtigte man zur Bestimmung des Ziels einer Pseudonymisierung nur diese Definition, müsste man davon ausgehen, dass der Teil der Daten, der nach der Abtrennung der Daten mit besonderem Identifikationspotential übrigbleibt, ohne Hinzuziehung der zusätzlichen Information nicht mehr personenbeziehbar sein dürfte. Der Personenbezug ist in der DSGVO in Art. 4 Nr. 1 nur positiv formuliert und nicht in Abgrenzung zu nicht mehr personenbeziehbaren Daten. Auch eine Definition anonymer Daten fehlt. Dies ist in einer gewissen Weise konsequent, da die DSGVO nur auf personenbeziehbare bzw. personenbezogene Daten anwendbar ist. In Erwägungsgrund Nr. 26 DSGVO finden sich aber

die entscheidenden Hinweise zur Abgrenzung von personenbeziehbaren zu nicht mehr personenbeziehbaren Daten. Letztere werden hier auch anonyme Daten genannt. Ein genauer Vergleich der Formulierungen zur Zielsetzung der Pseudonymisierung in Art. 4. Nr. 5 DSGVO mit den Formulierungen zur Grenze des Personenbezugs in Erwägungsgrund Nr. 26 DSGVO zeigt, dass die Formulierungen zur Zielsetzung der Pseudonymisierung hinsichtlich der nach Abtrennung übriggebliebenen Daten sogar noch absoluter und strenger sind, als es die in Erwägungsgrund Nr. 26 formulierte Abgrenzung zwischen noch personenbeziehbaren und nicht mehr personenbeziehbaren Daten eigentlich erfordern würde. Insoweit sind die Feststellungen in Erwägungsgrund Nr. 26 DSGVO bei der Auslegung der Kriterien in der Definition der Pseudonymisierung nach Art. 4 Nr. 5 DSGVO heranzuziehen.

Die Pseudonymisierung von Daten in der Forschung hat im deutschen Rechtsraum schon vor Anwendbarkeit der DSGVO weite Verbreitung gefunden. Beispielsweise wurde ein Leitfaden zum Datenschutz in der medizinischen Forschung mit detaillierten Vorgaben zur Pseudonymisierung von Gesundheitsdaten und Bioproben von der TMF erarbeitet und umfangreich mit den Datenschutzbehörden in Deutschland abgestimmt [4]. Grundlage hierfür war u. a. eine Definition im früheren Bundesdatenschutzgesetz. Weder sah die Definition vor, dass die nach der Abtrennung der direkt identifizierenden Informationen übriggebliebenen Daten ohne Hinzuziehung der abgetrennten Daten so gut wie anonym sein müssen, noch wurde dies in der Anwendung regelhaft so umgesetzt.

Vor diesem Hintergrund lag die Frage nach der Interpretation der Zielsetzung der Pseudonymisierung in Art. 4 Nr. 5 DSGVO mit der Einführung der DSGVO nahe. Die TMF hat hierzu eine umfangreiche gutachterliche Stellungnahme eingeholt, die zu dem Ergebnis kommt, dass für ein umfassendes Verständnis des Konzepts der Pseudonymisierung in der DSGVO die Heranziehung der Definition in Art. 4 Nr. 5 alleine nicht ausreichend ist, sondern vielmehr alle Verwendungen bzw. Verweise auf das Konzept in der DSGVO berücksichtigt werden müssen [1, S. 174ff]. Das von der TMF beauftragte Gutachten kommt zu dem Schluss, dass das Ergebnis einer Pseudonymisierung durchaus sein kann, dass ein Teil der Daten nach einer Abtrennung bestimmter Informationen nach den Kriterien von Erwägungsgrund Nr. 26 keinen Personenbezug mehr aufweist, dass aber in aller Regel das Ergebnis ein weiterhin personenbezogener Teildatensatz ist, der lediglich im Sinne einer Datenminimierung so reduziert ist, dass sich in diesem Teildatensatz keine Daten mehr befinden, die für die angestrebten Ziel der Verarbeitung nicht erforderlich sind. Das Gutachten löst hier das Spannungsverhältnis zwischen dem Erfordernis einerseits so gut wie anonymer und andererseits personenbeziehbarer Daten dahingehend auf, dass im Ergebnis zwei Arten pseudonymer Daten als möglich angesehen werden. Um die Argumentation des Gutachtens zu verstehen, ist es notwendig, sich mit der Verortung pseudonymer Daten zwischen Personenbezug und Anonymität zu befassen.

Zur Verortung pseudonymer Daten zwischen Personenbezug und Anonymität

Die Leitlinie verpasst überwiegend die Gelegenheit, den Begriff der Pseudonymisierung zu schärfen und so bestehende Unsicherheiten bei den Rechtsanwendern zu beseitigen. Vor dem Hintergrund der Bedeutsamkeit der Pseudonymisierung im Regelwerk der DSGVO, wäre eine

klare Positionierung wünschenswert. Insbesondere vermisst die TMF eine Auseinandersetzung mit den Folgen des relativen Begriffs von Personenbezug in Mehrpersonenkonstellationen.

Ausgehend von der Begriffsbestimmung des Art. 4 Nr. 5 DSGVO ist eine Lesart der Pseudonymisierung möglich, nach der nach der Trennung der Zusatzinformationen von den übrigen Daten, diese übrigen und dann pseudonymisierten Daten nicht mehr personenbeziehbar sein sollen. Die Leitlinie verwendet den Begriff der Zusatzinformationen unterscheidungslos sowohl für solche Informationen, die beim Vorgang der Pseudonymisierung abgetrennt wurden als auch Informationen, die unabhängig hiervon sowieso an anderer Stelle existieren (als Beispiel nennt die Leitlinie Social-Media-Posts oder Posts in Foren, Abs. 21). Die Frage der Zuordenbarkeit letzterer Informationen ist keine Besonderheit der Pseudonymisierung, sondern ein allgemeines Kriterium für Personenbezug. Und so verweist auch die Definition der Pseudonymisierung in Art. 4 Nr. 5 DSGVO hinsichtlich der für die Identifizierbarkeit der nach Abtrennung übriggebliebenen Daten relevanten Informationen eindeutig nur auf die im Rahmen der Pseudonymisierung abgetrennten Informationen, nicht aber auf beliebige Informationen an anderer Stelle.

Tatsächlich wird der Begriff der Pseudonymisierung in der DSGVO an vielen Stellen und in unterschiedlichen Kontexten verwendet. In Erwägungsgrund Nr. 26 DSGVO wird festgestellt:

> *[…] Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person. […]*

Insofern entsteht ein gewisses Spannungsverhältnis zwischen der Beschreibung pseudonymer Daten ohne Zugriff auf die abgetrennten Informationen als so gut wie anonym in der Definition in Art. 4 Nr. 5 DSGVO und der Feststellung, dass pseudonyme Daten immer als personenbeziehbar gelten sollen in Erwägungsgrund Nr. 26 DSGVO. Entscheidend ist hier offensichtlich die Anrechenbarkeit eines Zugriffs auf die abgetrennten Informationen für den Verantwortlichen, der die nach Abtrennung übriggebliebenen pseudonymen Daten verarbeitet. Hier sind Konstellationen denkbar und in bestimmten Fällen auch umgesetzt, in denen der Verantwortliche, der die pseudonymen Restdaten verarbeitet, keinen Zugriff auf die abgetrennten Informationen hat. Für solche Konstellationen muss nach der Rechtsprechung des EuGH davon ausgegangen werden, dass diese Verarbeitung keine Verarbeitung personenbeziehbarer Daten mehr darstellt [2; 3].

Vor diesem Hintergrund lässt sich das Spannungsverhältnis zwischen der Forderung nach einerseits nicht mehr personenbeziehbaren Daten sowie andererseits der Feststellung, dass immer von einer Personenbeziehbarkeit ausgegangen werden soll, nur auflösen, in dem man zwei unterschiedliche mögliche Ergebnisse einer Pseudonymisierung berücksichtigt. In dem ersten Fall kann die Verarbeitung der pseudonymen Restdaten tatsächlich als Verarbeitung anonymer Daten gewertet werden, in dem zweiten Fall ist weiterhin von der Verarbeitung personenbeziehbarer Daten auszugehen. In dem zweiten Fall ist die Pseudonymisierung lediglich als datenminimierende Maßnahme zu verstehen, wie es auch die die Nennung in Art. 89 Abs. 1 DSGVO nahelegt:

> *[…] Those safeguards shall ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimisation. Those measures may include pseudonymisation provided that those purposes can be fulfilled in that manner. […]*

Bei der Frage, ob die Verarbeitung der nach der Abtrennung direkt identifizierender Informationen übrigbleibenden pseudonymen als Verarbeitung anonymer Daten gewertet werden kann ist einerseits die Zugreifbarkeit auf die abgetrennten Informationen und andererseits die Zugreifbarkeit weiterer, die Identifizierung ermöglichender Informationen, zu berücksichtigen. Die Pseudonymisierung selbst hebt nur auf die erste Möglichkeit ab, für die Feststellung einer Anonymität ist aber immer eine umfassende Bewertung notwendig, wie es richtigerweise auch in der Leitlinie in Abs. 22 festgestellt wird:

> *[…] Even if all additional information retained by the pseudonymising controller has been erased, the pseudonymised data becomes anonymous only if the conditions for anonymity are met.*

Vor dem Hintergrund dieser Analyse ist nun die Verortung der Leitlinie und eine ausführliche Kritik derselben vorzunehmen.

Die Verortung pseudonymer Daten nach der Leitlinie

Wie geht nun die Leitlinie mit diesen zwei möglichen Ergebnissen einer Pseudonymisierung um? Weder ist erkennbar, dass das Gutachten [1, S. 174ff] selbst Berücksichtigung in der Leitlinie gefunden hätte, noch dass in der Leitlinie einer vergleichbar gründlichen Herangehensweise bei der Interpretation des Konzepts der Pseudonymisierung gefolgt worden wäre. Anstatt zwei mögliche und sich deutlich unterscheidende Zielsetzungen der Pseudonymisierung zu unterscheiden, folgt die Leitlinie eher einem unscharfen Mittelweg, in dem sie suggeriert, dass man mit einer Pseudonymisierung den nach der Abtrennung der offensichtlich identifizierenden Daten übriggebliebenen Teil möglichst nahe an einen anonymen Datensatz heranführen müsste, ohne aber Anonymität notwendigerweise auch zu erreichen. Als Kriterium für die Effektivität der Pseudonymisierung wird lediglich auf eine ausreichende Risikominimierung hingewiesen, die zum einen von Vorgaben des Gesetzgebers bzw. den Kriterien zur Anwendbarkeit einer bestimmten Rechtsgrundlage für die Verarbeitung abhängt (Abs. 24):

> *Union or Member State law may require pseudonymisation of personal data for the processing of personal data in specific situations, e.g. when providing for a legal basis under Art. 6(1)(c) or (e) GDPR in accordance with Art. 6(3) GDPR, or as a further condition in accordance with Art. 9(4) GDPR. In such cases, the law may also lay down specific requirements the pseudonymisation process or output has to meet, or the objectives it should achieve.*

Hier fällt auf, dass die Leitlinie es auch dem nationalen Gesetzgeber überlassen würde, nicht nur vorzuschreiben, wann eine Pseudonymisierung anzuwenden ist, sondern darüber hinaus auch, wann eine Pseudonymisierung eine ausreichende Risikoreduzierung erreicht. Letzteres würde aber bedeuten, wenn das Maß der Risikoreduzierung ein Kriterium dafür ist, ob eine Pseudonymisierung effektiv ist oder nicht, dass der nationale Gesetzgeber Kriterien für das Erreichen einer Pseudonymisierung vorgeben und damit auch Einfluss auf die Definition der Pseudonymisierung nehmen könnte. Dies kann offensichtlich vom europäischen Gesetzgeber nicht gewollt sein.

Zum anderen überlässt die Leitlinie eine ausreichende Risikoreduktion dem Verantwortlichen selbst (Abs. 25):

> *When such specific mandates for pseudonymisation are absent, controllers themselves may define the objectives that pseudonymisation should achieve. Those objectives may be connected with the processing they intend to perform themselves or with any subsequent processing of the pseudonymised data by recipients of those data.*

Diese Festlegung auf ein Kriterium einer ausreichenden Risikoreduktion als Ziel einer Pseudonymisierung übersieht, dass typischerweise eine Risikoanalyse immer das Ergebnis eines ganzen Bündels an technischen und organisatorischen Maßnahmen im Zusammenspiel betrachtet und nicht alleine auf eine Maßnahme schaut. Dies wird so auch in Art. 35 Abs. 7 lit. d zu den notwendigen Inhalten einer Datenschutz-Folgenabschätzung sehr deutlich formuliert:

> *[The assessment shall contain at least:] the measures envisaged to address the risks, including safeguards, security measures and mechanisms to ensure the protection of personal data and to demonstrate compliance with this Regulation taking into account the rights and legitimate interests of data subjects and other persons concerned.*

Mangels klarer Benennung unterschiedlicher Ergebnisse einer Pseudonymisierung von der vollständigen Anonymität einerseits bis zu einer reinen Datenminimierung gemäß den Kriterien der Erforderlichkeit andererseits baut die Leitlinie hier ein Risikokriterium auf, welches die Pseudonymisierung als technische und organisatorische Maßnahme, deren Wirkung in aller Regel im Zusammenspiel mit anderen Maßnahmen abzuschätzen ist, überfrachtet und damit die Anwendbarkeit erschwert und die Nutzung z. B. für Forschungszwecke einschränkt.

Sehr deutlich wird dies bei der Behandlung von Identifizierungsrisiken in einem pseudonymen Datensatz. Variablen in einem solchen Datensatz, deren Ausprägungen in Kombination möglicherweise die Identifizierung einzelner Betroffener erlauben könnten – hier quasi-identifier genannt – sollen analysiert werden und eine Nachbehandlung erfordern. Dabei ist genau das Bestehen solcher Identifizierungsrisiken Definitionskern des Personenbezugs. Bestünden solche Risiken nicht, hätte man – keinen Zugriff auf die abgetrennten direkt identifizierenden Daten vorausgesetzt – anonyme Daten vor sich. Als mögliche quasi-identifier werden genannt (Abs. 101):

> *[…] attributes contained in the data that reveal information about the physical, physiological, genetic, mental, economic, cultural or social identity of the data subject. If a combination of those attributes are sufficient to attribute at least part of the pseudonymised data to data subjects, then they are called quasi-identifiers. Demographic data are prime examples of such attributes: age, gender, languages spoken, marital or family status, profession, income. […]*

Und für die Nachbehandlung wird dann in Abs. 101 vorgeschlagen:

> *The most direct way to prevent attribution based on quasi-identifiers is their removal. A second approach lies in their modification by generalisation and randomisation.*

Solche Maßnahmen aus dem Bereich der Anonymisierung, die entsprechend unumkehrbar sind, führen aber gerade dazu, die Datenqualität eines Datensatzes erheblich zu beeinträchtigen. Die TMF beschäftigt sich seit über 10 Jahren intensiv mit der Anonymisierung

von Forschungsdaten und hat hierzu ein Weiterbildungsangebot ausgearbeitet, was seit vielen Jahren umfangreich genutzt wird. Aus dieser intensiven Befassung heraus kann aber auch abgeleitet werden, dass solche Maßnahmen allenfalls dann – und auch dann nur in sehr bestimmten Fällen – anwendbar sind, wenn die Fragestellung, die mit den Daten beantwortet werden soll, vollständig bekannt ist.

Wie Erwägungsgrund Nr. 33 DSGVO richtig reflektiert, lässt sich aber gerade in der Forschung oftmals zum Zeitpunkt der Erhebung der Daten noch nicht vollständig sagen, für welche wichtigen Fragestellungen diese zu einem späteren Zeitpunkt noch verwendet können. Diese notwendige Berücksichtigung künftiger Fragestellungen führt dazu, dass keine Variable, die später noch einmal wichtig werden könnte, zum jetzigen Zeitpunkt schon in unumkehrbarer Weise vergröbert oder gar gelöscht werden dürfte. Wenn z. B. das Alter der Betroffenen für eine spätere Auswertung als Gruppierungskriterium (unabhängige Variable) genutzt werden soll, würde ein zufälliges Verrauschen der Werte die Genauigkeit künftiger Auswertungen direkt negativ beeinflussen. Eine Generalisierung in Richtung der späteren Gruppierungsgrenzen wäre aber vielleicht unschädlich. Würde hingegen das Alter zweier Patientengruppen als auszuwertende Größe (abhängige Variable) angesehen, wäre ein geringes Verrauschen möglicherweise weniger kritisch als eine grobe Generalisierung.

Hier droht die Verschiebung der Zielsetzung der Pseudonymisierung in Richtung einer Anonymisierung die Maßnahme gerade für die Forschung, in der diese Maßnahme heute noch weit verbreitet und breit akzeptiert ist, unattraktiv zu machen. Nicht auszuschließen, dass in Folge einer solchen Umdeutung selbst die basale und in der Praxis sehr hilfreiche Maßnahme der Abtrennung direkt identifizierender Daten seltener als bisher Anwendung findet.

Dort, wo im nationalen oder auch europäischen Recht heute schon eine Pseudonymisierung vorgeschrieben ist oder auch künftig beispielswiese im Kontext des European Health Data Space (EHDS) vorgeschrieben wird, könnten sehr viele Anwendungsfälle aus der wissenschaftlichen Forschung durch eine solche Verortung der Pseudonymisierung ausgeschlossen werden.

Hinsichtlich des Ziels einer bestimmten Risikoreduktion, die die Leitlinie alleine der Pseudonymisierung aufbürdet, verdeckt sie damit leider auch den Blick darauf, dass die Pseudonymisierung natürlich um weitere Maßnahmen ergänzt werden kann, um bestimmte Ziele zu erreichen. Insbesondere können Maßnahmen, die in Ziff. 3.1.3 der Leitlinie benannt werden, solche sinnvollen Ergänzungen darstellen. Eine solche Möglichkeit wird zwar in Abs. 96 angedeutet, letztlich aber nicht für eine klare Abgrenzung der Pseudonymisierung von anderen Maßnahmen genutzt:

> *Controllers can benefit from a potential trade-off: the smaller the pseudonymisation domain and the more restrictive the access to pseudonymised data and other relevant information sources within the pseudonymisation domain, the less need there is in general, considering the remaining circumstances, to modify the original data.*

Positiv ist hervorzuheben, dass die Leitlinie explizit darauf hinweist, dass die Pseudonymisierung die Einhaltung von Voraussetzungen einer datenschutzgerechten Verarbeitung unterstützen kann (Abs. 30):

> *The effectiveness of the implementation of pseudonymisation determines the extent of the reduction of risks for the data subjects and the benefits the controllers may derive from it, including the fulfilment of data-protection obligations according to Art. 24, 25 and 32 GDPR […]*

Nur muss eben deutlich gesagt werden, dass diese Aufgabe der Einhaltung eines ausreichenden Schutzes der Rechte und Freiheiten der betroffenen Personen in aller Regel nicht der Pseudonymisierung alleine, sondern dem Zusammenspiel unterschiedlicher technischer und organisatorischer Maßnahmen zukommt.

Diese Vorgabe einer wie auch immer bemessenen Risikoreduktion ist weder für die Anwender einer Pseudonymisierung hilfreich, die damit im Sinne der Datenminimierung handeln wollen, ohne aber Anwendungsfälle oder die Datenqualität einzuschränken, noch für die Anwender, die im Ergebnis einer Pseudonymisierung möglicherweise tatsächlich im rechtlichen Sinne nicht mehr personenbeziehbare Daten ansteuern.

Fehlende Anerkennung des Konzepts eines relativen Personenbezugs

Da die Pseudonymisierung gemäß Definition in der DSGVO ein Vorgang der Aufteilung von Daten mit unterschiedlichen Eigenschaften hinsichtlich der Identifizierbarkeit und deren getrennter Weiterverarbeitung ist, stellt sich hier in besonderer Weise die schon seit langem kontrovers diskutierte Frage danach, ob man das Konzept des Personenbezugs, welches der DSGVO zugrunde liegt, als ein relatives oder absolutes verstehen muss. Der relative Ansatz würde dazu führen, dass die Verarbeitung von Daten, von denen alle personenbeziehbaren Informationen abgetrennt wurden, in einer Weise, dass kein Zugriff auf die abgetrennten Daten mehr besteht, grundsätzlich auch als Verarbeitung anonymer Daten verstanden werden könnte, der unabhängig von den Vorgaben der DSGVO durchführbar wäre. Der absolute Ansatz hingegen schaut nicht darauf, welche Möglichkeiten der Verantwortliche hat, eine Person anhand der ihm zugänglichen Daten zu identifizieren, sondern vielmehr darauf, ob es an irgendeiner Stelle noch Zuordnungsdaten gibt, die eine Identifizierung ermöglichen würden. Pseudonyme Daten, zu denen es ja per Definition Zuordnungsdaten gibt, wären demnach immer personenbeziehbar. Da der EuGH zu dieser Frage schon Stellung bezogen hat und zu dem Ergebnis gekommen ist, dass der DSGVO ein relatives Verständnis zugrunde liegt [2; 3], ist es bedauerlich, dass sich die Leitlinie des EDSA hierzu nicht eindeutig positioniert.[2] Vielmehr gibt es in der Leitlinie viele Formulierungen, die ein absolutes Verständnis nahelegen und damit den Möglichkeitsraum der Pseudonymisierung unnötig einengen (z. B. Abs. 22):

> *Pseudonymised data, which could be attributed to a natural person by the use of additional information, is to be considered information on an identifiable natural person, and is therefore personal. This statement also holds true if pseudonymised data and additional information are not in the hands of the same person. […]*

[2] Auch die Schlussanträge des Generalanwalts in der Rechtssache C-413/23 P weisen in dieselbe Richtung (https://curia.europa.eu/juris/document/document.jsf?pageIndex=0&docid=295078&doclang=EN&text=&cid=5888244).

Das Anerkennen des Konzepts des relativen Personenbezugs in Übereinstimmung mit der Rechtsprechung des EuGH wäre wünschenswert und hilfreich für viele Anwender.

Literatur

1. Dierks, C., Roßnagel, A., *Sekundärnutzung von Sozial- und Gesundheitsdaten – Rechtliche Rahmenbedingungen*. 2019, Medizinisch Wissenschaftliche Verlagsgesellschaft, Berlin, https://mwv-open.de/site/books/10.32745/9783954665181/ (Abruf: 2025-03-10).

2. EuGH *Urteil des Gerichtshofs (Zweite Kammer) vom 19. Oktober 2016. Patrick Breyer gegen Bundesrepublik Deutschland. Aktenzeichen C-582/14*. 2016. Europäischer Gerichtshof, http://curia.europa.eu/juris/document/document.jsf?docid=184668&doclang=DE (Abruf: 2025-03-10).

3. EuGH *Urteil des Europäischen Gerichtshofes (Dritte Kammer) vom 9. November 2023. Gesamtverband Autoteile-Handel e. V. gegen Scania CV AB. Aktenzeichen C-319/22*. 2023. Europäischer Gerichtshof, https://eur-lex.europa.eu/legal-content/de/TXT/?uri=CELEX:62022CJ0319 (Abruf: 2025-03-10).

4. Pommerening, K., Drepper, J., Helbing, K., Ganslandt, T., *Leitfaden zum Datenschutz in medizinischen Forschungsprojekten – Generische Lösungen der TMF 2.0*. 2014, Medizinisch Wissenschaftliche Verlagsgesellschaft, Berlin, https://www.mwv-open.de/site/books/10.32745/9783954662951/ (Abruf: 2025-03-10).

Abkürzungsverzeichnis

DSGVO	Verordnung des Europäischen Parlaments und des Rates zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG – Datenschutz-Grundverordnung (Verordnung 2016/679)
EDSA	Europäischer Datenschutzausschuss (https://edpb.europa.eu)
EG	Europäische Gemeinschaft
EHDS	European Health Data Space, Regulierung der Europäischen Kommission zu einem europäischen Gesundheitsdatenraum (https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en)
EuGH	Gerichtshof der Europäischen Gemeinschaften (http://curia.europa.eu)
GDPR	General Data Protection Regulation
lit	Littera / Buchstabe
TMF	TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. (www.tmf-ev.de)