Datum 23. März 2020

STELLUNGNAHME IM KONSULTATIONSVERFAHREN ZUM ENTWURF DES POSITIONSPAPIERS ZUR ANONYMISIERUNG UNTER DER DSGVO UNTER BERÜCKSICHTIGUNG DER TK-BRANCHE DES BUNDESBEAUFTRAGTEN FÜR DEN DATENSCHUTZ UND DIE INFORMATIONSFREIHEIT

Hinweis: Wir sind mit der Veröffentlichung unserer Stellungnahme einverstanden.

Inhaltsverzeichnis

Einleitung

Der Wunsch nach der freien Nutzung möglichst aussagekräftiger Daten ist weit verbreitet. Sowohl öffentliche als auch private Einrichtungen aus Forschung und Wirtschaft bedürfen zur Evaluierung oder Weiterentwicklung der eigenen Produkte und Dienstleistungen immer mehr Daten, für Zwecke, die über die ursprünglichen Zwecke bei der Erhebung dieser hinausgehen oder unabhängig von diesen genutzt werden sollen. Sofern vom Verantwortlichen einmal erhobene personenbezogene Daten (weiter)verwendet werden sollen, ist stets eine entsprechende datenschutzrechtliche Grundlage notwendig. Hier hilft – wenn überhaupt – i.d.R. nur die Einwilligung des Betroffenen weiter. Für viele dieser Ziele sind allerdings bereits Datensätze mit nicht-personenbezogenem, sondern abstraktem Gehalt ausreichend. In diesen Fällen wäre für die Erreichung dieser Ziele die Information bzgl. einer konkret zugeordneten Person folglich nicht notwendig. Hierfür kommen insbesondere anonymisierte Daten in Betracht, die für die Entwicklung von KI und Big Data Anwendungen, für Marktforschung, statistische Analysen, Produktinnovation, sowie für Open Data Initiativen genutzt und mit Kollaborationspartnern geteilt werden können.

Die Europäische Datenschutzgrundverordnung („**DSGVO**“) sieht im Prinzip der Datenminimierung in solchen Fällen sogar ausdrücklich vor, anonymisierte Daten zu verwenden. Eben diese Anonymisierung ist im Datenschutzrecht generell und in der DSGVO im Speziellen jedoch nicht ausreichend geregelt. Weder ist eindeutig definiert, wann personenbezogene Daten durch die zur Verfügung stehenden Anonymisierungstechniken ausreichend anonymisiert wurden, sodass die DSGVO keine Anwendung mehr findet und die Daten entsprechend „frei“ genutzt werden dürfen, noch ist die Rechtmäßigkeit der Anonymisierung hinreichend konkretisiert. Dadurch entsteht bei den Anwendern Unsicherheit, ob, wann und wie Anonymisierungstechniken wirksam eingesetzt werden können. Diese Unsicherheiten hemmen die Entwicklung und Nutzung entsprechender Anonymisierungstechniken, wodurch insbesondere der Fortschritt im Bereich Künstlicher Intelligenz und der Big Data Nutzung aufgehalten wird. Eine Entwicklung, die erhebliche negative Auswirkungen auf die digitale Wirtschaft Deutschlands und der Europäischen Union haben wird, da in diesem Bereich bereits jetzt amerikanische und chinesische Unternehmen auf dem Vormarsch sind. Nicht ohne Grund handelt es sich daher um eine Entwicklung, welcher die Europäische Kommission aktiv entgegenwirken möchte. Beispielsweise durch ihre 2020 veröffentlichte KI- und Datenstrategie, durch die die EU zu einem globalen Vorreiter im Bereich der Digitalisierung werden und durch die ermöglicht werden soll, dass europäische Organisationen – unter Einhaltung europäischer Werte und Rechte wie dem Datenschutz – die stetig wachsenden Menge an Daten vermehrt nutzen und mit dieser innovieren.

Für die Herbeiführung entsprechender Rechtssicherheit ist es daher zwingend notwendig,

(i) einen einheitlichen rechtlichen Begriff „ausreichender Anonymisierung“ zu definieren;

(ii) festzustellen, ob eine Anonymisierung eine Datenverarbeitung darstellt, die in der Folge einer Rechtsgrundlage bedarf;

(iii) sofern die Anonymisierung eine rechtfertigungsbedürftige Datenverarbeitung darstellt, diese zum Schutz der betroffenen Personen im Gegensatz zu anderen Verarbeitungen datenschutzrechtlich zu privilegieren, oder jedenfalls

(iv) festzustellen, dass für eine ausreichende Anonymisierung stets eine Rechtsgrundlage nach Art. 6 Abs. 4 DSGVO (Weiterverarbeitung) oder Art. 6 Abs. 1 lit. f DSGVO (berechtigte Interessen des Verantwortlichen) vorliegt und

(v) festzustellen, dass, sofern eine Datenverarbeitung vorliegt, Verantwortliche zumindest im Hinblick auf die weiteren, aus der Natur der Datenverarbeitung resultierenden Pflichten nach der DSGVO privilegiert sind. Dies betrifft insbesondere die Informationspflichten.

Auf all diese Punkte geht das Positionspapier bislang nicht oder nur unzureichend ein.

I. Executive Summary

- Mit neuartigen Anonymisierungstechniken, bspw. der Synthetisierung, ist mittlerweile eine **absolute Anonymisierung** personenbezogener Daten möglich. Aufgrund der fortschreitenden technischen Entwicklung ist es für Anwender von Anonymisierungstechniken sowie aus Gründen der Rechtssicherheit insgesamt dennoch erforderlich, **Anforderungen zu definieren, wann eine Anonymisierungstechnik hinreichend anonymisierte Daten erzeugt**.

- Demgegenüber ist **nach der DSGVO bereits eine lediglich faktische Anonymisierung ausreichend**, d.h. eine De-Anonymisierung muss nicht 100% ausgeschlossen sein. Vielmehr ist es ausreichend, wenn das Restrisiko einer Identifizierung des Betroffenen soweit wie möglich ausgeschlossen wird und unter praktischen Erwägungen gewährleistet ist, dass (i) es nicht mehr möglich ist, eine bestimmte Person aus einem Datenbestand herauszugreifen, (ii) die eine Person betreffenden Datensätze miteinander zu verknüpfen oder (iii) durch Inferenz Informationen aus einem solchen Datenbestand über eine Person abzuleiten.

- **Die Anonymisierung stellt keine Datenverarbeitung im Sinne der DSGVO dar**:

  - **Unter Berücksichtigung des Schutzzwecks der DSGVO** kann eine Anonymisierung nicht unter den Begriff „Verarbeitung“ subsumiert werden. Denn im Gegensatz zu allen anderen Verarbeitungstätigkeiten nach der DSGVO ist **bei der Anonymisierung nicht die Verarbeitungstätigkeit an sich, sondern allein deren Ergebnis**, nämlich das Nichtvorhandensein von personenbezogenen Daten, entscheidend („**Anonymisierungsverarbeitung**“).

  - Entsprechend ist die Anonymisierungsverarbeitung eine tatsächliche **Umsetzung des Grundsatzes der Datensparsamkeit und Datenminimierung** und als solche keine klassische „Verarbeitung“ im Sinne der DSGVO.

  - Bereits **nach alter Rechtslage** war nach Ansicht des Bundesbeauftragen für den Datenschutz und die Informationsfreiheit („BfDI“) die **Anonymisierung keine Datenverarbeitung**.

- Hilfsweise ist die Anonymisierungsverarbeitung zur Umsetzung und Förderung des Prinzips der Datensparsamkeit **durch Entbehrlichkeit einer Rechtsgrundlage zu privilegieren**.

- Sollte jedoch, weiter hilfsweise, das Erfordernis einer Rechtsgrundlage bestehen, ist die **Anonymisierungsverarbeitung standardmäßig nach denselben rechtlichen Kriterien zu rechtfertigen**, denn sie erzeugt immer dasselbe Ergebnis (anonyme Daten):

    - Eine **Anonymisierung ist nach Art. 6 Abs. 4 DSGVO zulässig**, sofern es sich um die Anonymisierung **bereits vorhandener personenbezogener Daten** handelt, die zunächst **zu anderen Zwecken erhoben** wurden.

    - Sofern die **personenbezogenen Daten ausschließlich zum Zweck der Anonymisierung erhoben** werden, ist eine **Rechtfertigung nach Art. 6 Abs. 1 lit. f DSGVO** gegeben.

    - Jede darauffolgende **Nutzung der bereits anonymisierten Daten** ist – mangels Geltung der DSGVO – **ohne weitere datenschutzrechtliche Grundlage** möglich.

- **Sofern eine Anonymisierungsverarbeitung eine Datenverarbeitung im Sinne der DSGVO darstellt**, muss der Verantwortliche die Betroffenen über diese Verarbeitung **entsprechend informieren**. Unter Berücksichtigung des Ergebnisses der Verarbeitung (anonyme Daten) ist **hierbei ein allgemeiner Hinweis** darauf, dass personenbezogene Daten ggf. (i) einem Anonymisierungsverfahren unterzogen werden und (ii) diese im Anschluss nicht mehr auf eine natürliche Person zurückzuführen sind, **ausreichend**. Weitergehende Informationen hinsichtlich der anschließenden Verwendung der anonymen Daten sind hingegen nicht mehr erforderlich, da diese nicht mehr der DSGVO unterfallen.

II. Tatsächliche Anforderungen an die Anonymisierung

Im Zusammenhang mit der Anonymisierung von personenbezogenen Daten entstehen insbesondere Reibungspunkte hinsichtlich:

- der Erreichung eines **hinreichenden Maßes an Anonymität** der Daten durch Verallgemeinerung, Löschung, Verfälschung, Hinzufügung der vorhandenen Informationen oder Synthetisierung (im Folgenden zusammenfassend als „**Anonymisierungstechniken**“ bezeichnet) und

- der **Erhaltung des erforderlichen Maßes an (statistischer) Aussagekraft**, die durch das jeweilige Anonymisierungsverfahren ggf. verloren gehen kann.

Aufgrund der fortschreitenden technischen Entwicklung und Verfügbarkeit von zusätzlichen Daten ist eine De-Anonymisierung heutzutage immer einfacher möglich (z.B. durch den Abgleich vermeintlich anonymer Daten mit Informationen aus anderen Datenbanken).[1] Daher müssen mittlerweile hohe Anforderungen an die Wirksamkeit der Anonymisierung gestellt werden. In vielen Fällen wird anstatt einer Anonymisierung tatsächlich eine Pseudonymisierung erreicht, die jedoch unter Zuhilfenahme zusätzlicher Daten wieder rückgängig gemacht werden kann. Rechtlich sind die Anforderungen, die an diesen „Aufwand“ gestellt werden, umstritten. Für die Anwender von Anonymisierungstechniken sowie für die Rechtssicherheit ist es daher von großer Bedeutung, die entsprechenden Anforderungen festzusetzen und zu definieren, wann eine Anonymisierungstechnik hinreichend anonymisierte Daten erzeugt.

1. „Die Anonymisierung“

Die DSGVO definiert „Anonymisierung“ nicht – im Gegensatz zur Pseudonymisierung in Art. 4 Nr. 5 DSGVO. Lediglich in Erwägungsgrund 26 heißt es hierzu:

> „*Die Grundsätze des Datenschutzes sollten daher nicht für anonyme Informationen gelten, d.h. für Informationen, die sich nicht auf eine identifizierte oder identifizierbare natürliche Person beziehen (*Alternative 1*), oder personenbezogene Daten, die in einer Weise anonymisiert worden sind, dass die betroffene Person nicht oder nicht mehr identifiziert werden kann (*Alternative 2*). Diese Verordnung betrifft somit nicht die Verarbeitung solcher anonymer Daten, auch für statistische oder für Forschungszwecke.“*

[1] *Ernst*, in: Paal/Pauly, DS-GVO BDSG, 2. Aufl. 2018, Art. 4 Rn. 50.

(Hervorhebungen durch den Verfasser).

Damit ist bei der Anonymisierung ein Mehr verlangt gegenüber der Pseudonymisierung, bei der zur Identifikation der dahinterstehenden Person, die Hinzuziehung zusätzlicher, aber getrennt aufbewahrter Informationen genügt.[2] Anonymisierte Daten hingegen weisen keinen Personenbezug mehr auf. Auf sie ist folglich auch die DSGVO nicht (mehr) anwendbar.[3] Dies gilt sowohl für die absolute als auch für die faktische Anonymisierung.[4]

Eine **absolute Anonymisierung** liegt vor, wenn unabhängig von möglichem Zusatzwissen eines Dritten für niemanden eine De-Anonymisierung mehr möglich ist.[5] Eine solche absolute Anonymisierung ist nur mit sehr wenigen der derzeit verfügbaren Methoden zu erreichen. Sie ist datenschutzrechtlich allerdings gar nicht notwendig.[6] Dies folgt ebenfalls aus Erwägungsgrund 26 der DSGVO, der insoweit „lediglich" auf die technische und wirtschaftlich mögliche Wahrscheinlichkeit einer De-Anonymisierung abstellt:

> *„Um festzustellen, ob eine natürliche Person identifizierbar ist, sollten alle Mittel berücksichtigt werden, die von dem Verantwortlichen oder einer anderen Person nach allgemeinem Ermessen wahrscheinlich genutzt werden, um die natürliche Person direkt oder indirekt zu identifizieren, wie beispielsweise das Aussondern. Bei der Feststellung, ob Mittel nach allgemeinem Ermessen wahrscheinlich zur Identifizierung der natürlichen Person genutzt werden, sollten alle objektiven Faktoren, wie die Kosten der Identifizierung und der dafür erforderliche Zeitaufwand, herangezogen werden, wobei die zum Zeitpunkt der Verarbeitung verfügbare Technologie und technologische Entwicklungen zu berücksichtigen sind."*
>
> (Hervorhebungen durch den Verfasser).

Die DSGVO geht folglich ebenfalls neben der absoluten Anonymisierung (Alternative 1) davon aus, dass eine faktische Anonymisierung (Alternative 2) ausreichend ist. Die **faktische Anonymisierung** gewährleistet also nicht absolut und in jedem Fall den Verlust von Personenbezug. Es wird vielmehr unter praktischen Erwägungen gewährleistet, dass ein „Angreifer" auf diese Informationen – also ein Dritter, der weder der für die Verarbeitung Verantwortlicher noch Auftragsverarbeiter ist[7] – einen so hohen Aufwand betreiben müsste,

[2] *Ernst*, in: Paal/Pauly, DS-GVO BDSG, 2. Aufl. 2018, Art. 4 Rn. 48.
[3] Vgl. Erwägungsgrund 26 der DSGVO sowie *Ernst*, in: Paal/Pauly, DS-GVO BDSG, 2. Aufl. 2018, Art. 4 Rn. 49.
[4] *Ziebarth*, in Sydow: Europäische Datenschutzgrundverordnung, 2. Aufl. 2018, Art. 4 Rn. 32.
[5] *Ziebarth*, in Sydow: Europäische Datenschutzgrundverordnung, 2. Aufl. 2018, Art. 4 Rn. 29.
[6] *Ziebarth*, in Sydow: Europäische Datenschutzgrundverordnung, 2. Aufl. 2018, Art. 4 Rn. 29.
[7] Artikel-29-Datenschutzgruppe, WP 216, S. 14.

dass dieser nicht mehr im Verhältnis zum potentiellen Nutzen stehen würde und damit wahrscheinlich von einem Angriffsversuch absehen würde.[8]

Die Artikel-29-Datenschutzgruppe hat in ihrem Working Paper 2016 einen Maßstab für eine **wirksame faktische Anonymisierung** definiert. Hiernach darf es nach einer Anonymisierung personenbezogener Daten, keiner Partei möglich sein,

- eine Person aus einem Datenbestand **herauszugreifen**,

- eine Person betreffende Datensätze **miteinander zu verknüpfen**, bspw. eine Verbindung zwischen zwei Datensätzen eines Datenbestands (oder zwischen zwei unabhängigen Datenbeständen) herzustellen, oder

- durch **Inferenz,** Informationen aus einem solchen Datenbestand über eine Person abzuleiten.[9] Hierbei muss weiter unterschieden werden zwischen personenbezogener Inferenz, d.h. Information die spezifisch zu einer identifizierbaren Person abgeleitet wird, egal ob dies deterministisch oder nur mit gewissen Wahrscheinlichkeiten geschlossen werden kann, und statistischer Inferenz, d.h. Wahrscheinlichkeiten, die sich aus der Gesamtpopulation ergeben, und die sich ebenso ergeben würden, wenn die Person, über die die Inferenz durchgeführt wird, in den Daten gar nicht selbst abgebildet ist.

Sofern die zuvor genannten Voraussetzungen gegeben sind, könnten Anonymisierungstechniken Garantien für den Schutz der Privatsphäre für Betroffene schaffen und Verantwortliche die Daten für andere Zwecke nutzen. Selbst in dem Fall, in dem ein Anonymisierungsverfahren eines der genannten Kriterien nicht erfüllt ist, müsste hierauf nicht verzichtet werden. Vielmehr sollte dann eine Evaluierung der hinsichtlich einer Identifizierung bestehenden Risiken für den Betroffenen vorgenommen werden.[10]

[8] vgl. Stellungnahme des LDI NRW, von April 2017, https://www.ldi.nrw.de/mainmenu_Datenschutz/submenu_Technik/Inhalt/TechnikundOrganisation/Inhalt/Anonymitaet-in-Zeiten-von-Big-Data/Anonymitaet-in-Zeiten-von-Big-Data1.pdf, zuletzt aufgerufen am 20. März 2020.
[9] Artikel-29-Datenschutzgruppe, WP 216, S. 3, 11.
[10] Artikel-29-Datenschutzgruppe, WP 216, S. 29.

Letztlich geht es bei der Beurteilung von Anonymisierungstechniken also darum, das mit ihnen verbundene Restrisiko einer Identifizierung des Betroffenen zu berücksichtigen und soweit wie möglich auszuschließen.

Für diese Evaluierung bedarf es vor Einsatz einer Anonymisierungstechnik, einer entsprechenden Planung durch Prüfung der jeweiligen Stärken und Schwächen sowie Festlegung der Voraussetzungen und der jeweiligen Zielsetzung des Anonymisierungsverfahrens. Die Wahl dieser jeweils am besten geeigneten Lösung sollte auf der Grundlage einer Einzelfallbewertung erfolgen.[11]

Im Ergebnis ist durch die Wahl des bzw. der (ggf. kombinierten Verfahren) also eine wirksame Anonymisierung möglich. Der Maßstab hierfür ist keine absolute Anonymisierung, sondern eine faktische Anonymisierung, bei der das Restrisiko einer Identifizierung unter Berücksichtigung der Umstände des Einzelfalles soweit wie möglich ausgeschlossen ist.

2. Voraussetzungen und Verfahren

Wie dargestellt definiert die DSGVO nicht, wie Daten wirksam anonymisiert werden, sondern geht vom Ergebnis aus: die Anonymisierung muss unumkehrbar sein. Grundsätzlich können daher unterschiedliche Anonymisierungstechniken gewählt werden, solange sie im Ergebnis zu einer faktischen Anonymisierung führen.

Die Artikel-29-Datenschutzgruppe geht hierbei sogar soweit, dass nach dem Einsatz von Anonymisierung weder dem Verantwortlichen, noch einem Dritten möglich sein darf selbst unter Verwendung „*aller Mittel, die vernünftigerweise eingesetzt werden können*“, eine natürliche Person zu bestimmen.[12] In diesem Zusammenhang wird gerne aufgeführt, dass einige wenige charakteristische Merkmale häufig schon ausreichen würden, um eine natürliche Person (wieder) zu bestimmen. Hieraus wird häufig der Schluss gezogen, dass eine Anonymisierung grundsätzlich unmöglich sei. Das ist jedoch unzutreffend. Hinsichtlich der aktuell existierenden Anonymisierungstechniken bedeutet das zunächst, dass Verantwortliche dies bei der Auswahl ihrer Anonymisierungstechnik berücksichtigen und ein ihren Zielen entsprechendes Verfahren wählen müssten.[13] Mit der Wahl der entsprechenden Anonymisierungstechnik ist eine faktische Anonymisierung sodann durchaus möglich.

[11] Artikel-29-Datenschutzgruppe, WP 216, S. 4.
[12] Artikel-29-Datenschutzgruppe, WP 216, S. 6
[13] Siehe zuvor unter 1 sowie auch Artikel-29-Datenschutzgruppe, WP 216, S. 11.

Für eine erfolgreiche faktische Anonymisierung ist es notwendig, (i) direkte Identifikationsmerkmale zu eliminieren und (ii) eine Identifikation durch Verknüpfung mit korrelierendem Wissen, so weit wie möglich auszuschließen bzw. die Identifizierungswahrscheinlichkeit stark zu verringern:[14]

Erste notwendige Voraussetzung ist stets die Eliminierung der sog. expliziten bzw. „**direkten Identifikationsmerkmale“,** also Namen, Anschriften, Personenkennzeichen, Bankverbindungen oder Telefonnummern.

In der Regel reicht die Eliminierung direkter Identifikationsmerkmale jedoch nicht aus. Es verbleiben genügend weitere Merkmale, die eine Person zumindest **indirekt identifizieren** können, bspw. durch Verknüpfung mehrerer indirekter Informationen oder anderem korrelierendem Wissen (sog. „**Quasi-Identifikatoren**“). Allerdings ist davon auszugehen, dass alle Merkmale Quasi-Identifikatoren sind, da die Anonymisierung unabhängig davon über welches Zusatzwissen potentielle Angreifer verfügen und auf welche Merkmale sich dieses beziehen lässt, wirksam sein muss. Daher sind diese weiter zu entfremden, um den Personenbezug so weit wie möglich verschwinden zu lassen:

Für diese Anonymisierung gibt es vier übergeordnete Techniken:

- **Randomisierung** (zufällige Veränderung von Daten): hierbei wird eine zufällige Veränderung der Daten vorgenommen, d.h. hier werden die Merkmale nach vorab definierten zufälligen randomisierten Mustern verändert, etwa (i) indem die Werte eines Merkmals jeweils mit einer gewissen Wahrscheinlichkeit durch andere mögliche Merkmalsausprägungen ersetzt werden, (ii) ein zufälliger Wert zu den Werten addiert oder (iii) die Werte mit einem zufälligen Wert multipliziert werden. Ein Vorteil dieser Methode ist es, dass Zusammenhänge in den Daten weitestgehend erhalten bleiben, und die Abweichungen quantifiziert werden können. Da bei der Randomisierung des Weiteren keine exakten realen Zahlen in den anonymisierten Daten erhalten bleiben, können Schlüsse nur mehr mit gewissen Wahrscheinlichkeiten getroffen werden (man kann also erreichen, dass man aus den anonymisierten Daten nicht mehr ableiten kann, dass z.B. das Gehalt einer bestimmten Person 51.320€ wäre, sondern nur noch, dass es mit 98%

[14] vgl. Stellungnahme des LDI NRW, von April 2017, Stellungnahme des LDI NRW, von April 2017, https://www.ldi.nrw.de/mainmenu_Datenschutz/submenu_Technik/Inhalt/TechnikundOrganisation/Inhalt/Anonymitaet-in-Zeiten-von-Big-Data/Anonymitaet-in-Zeiten-von-Big-Data1.pdf, zuletzt aufgerufen am 20. März 2020.

Wahrscheinlichkeit zwischen 50.000 und 55.000€ liegt). Nachteil ist dagegen, dass aufgrund des Erhalts der Zusammenhänge zwischen den Merkmalen aller Personen alle drei oben genannten Angriffsarten, i.e. Herausgreifen, Verknüpfung und (personenbezogene) Inferenz möglich sind, und durch die Randomisierung nur erschwert und mit höherer Unsicherheit versehen werden. Diese Erschwernis bzw. Unsicherheit muss im Einzelfall evaluiert werden, um die Risiken für die Datensubjekte abzuschätzen.

- **Permutation** (zufällige Vertauschung von Daten): hierbei werden die Werte des Merkmals untereinander vertauscht, und somit die direkte Verbindung zwischen Daten und betroffener Person entfernt.[15] Vorteil dieser Methode ist, dass alle Merkmale für sich betrachtet exakt erhalten bleiben und man daher exakte Statistiken auf der Ebene einzelner Merkmale berechnen kann, etwa die genaue Anzahl von Einwohnern jedes Bezirks, das exakte Durchschnittseinkommen, etc. Allerdings gehen alle Informationen über Zusammenhänge verloren, man kann etwa kein Durchschnittseinkommen pro Bezirk berechnen. Von den oben genannten Angriffsszenarien ist daher Verknüpfung grundsätzlich ausgeschlossen, da die Zeilen des anonymisierten Datensatzes nicht mehr einzelnen Personen zugeordnet werden können Es können jedoch einzelne Werte herausgegriffen werden und personenbezogene Inferenz betrieben werden, da alle exakten Werte in den Daten erhalten bleiben (und umgekehrt alle Werte in den anonymisierten Daten exakt echten Datenpunkten entsprechen). Im Beispiel von Gehaltsdaten kann dies z.B. bedeuten, dass ein Angreifer, der durch externe Information weiß, wer die Person mit dem höchsten Gehalt ist, im anonymisierten Datensatz das höchste Gehalt suchen und dadurch das exakte Gehalt dieser Person erfahren.

- **Generalisierung** (insbesondere Aggregation): bei dieser Technik werden genaue Werte durch ungenauere Werte ersetzt, z.B. indem Daten zusammengefasst werden (Bsp.: Alter 25: zu Alter 20-30). Aus dem Gruppensatz kann in der Folge nicht mehr festgestellt werden, welchen genauen Wert eine Person innerhalb des groben Wertebereichs hat, was es weiter schwieriger macht, Personen durch Aussondern zu re-identifizieren. Im Extremfall können durch Aussondern keine einzelnen Individuen mehr isoliert werden, sondern nur noch Gruppen von (auf der vergröberten Granularität) identischen Personen – in diesem Fall hat man K-anonymität erreicht, wobei K die Größe der kleinsten Personengruppe ist, die

[15] Artikel-29-Datenschutzgruppe, WP 216, S. 14; *Hansen*, in: Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht, Art. 4 Rn. 54;

ausgefiltert werden kann. Da auch hier sowohl die Werte selbst (wenngleich vergröbert) ohne Fehler erhalten bleiben, und auch Zusammenhänge zwischen Merkmalen erhalten bleiben, sind potentiell alle drei Angriffsszenarien möglich und die jeweiligen Risiken müssen abgeschätzt und mitigiert werden. Herausgreifen und Verknüpfung sind im Allgemeinen nicht mehr in Form einer 1:1-Zuordnung möglich, aber eine Person kann einer Gruppe von in den anonymisierten Daten zugeordnet werden. Folglich sind Rückschlüsse auf die dahinterstehende natürliche Person zwar meist nur noch mit Unsicherheit, aber trotzdem mit klarem Personenbezug möglich. So kann man etwa herausfinden, dass die Person, die man herausgreifen möchte, einer von K (z.B. 10) Einträgen in den Daten entsprechen könnte, und kann die Information, die man über diese Person lernen möchte (z.B. Gehalt), auf die entsprechenden 10 Werte eingrenzen und spezifisch für diese Person schließen, dass sie mit jeweils 10% Wahrscheinlichkeit eines dieser 10 Gehälter haben muss.

- **Datensynthetisierung** (Erstellung komplett neuer, synthetischer Daten): Hierbei werden die echten Daten vollständig verworfen und durch neue, zufallsgenerierte Werte ersetzt. Die statistischen Verteilungen, nach denen die neuen Daten generiert werden, werden – meist mittels maschinellem Lernen – aus den echten Daten geschätzt, damit die synthetischen Daten den echten Daten statistisch möglichst ähnlich sind. Sowohl alle individuellen Werte aller Merkmale als auch die Zusammenhänge zwischen Merkmalen sind künstlich anhand des aus den Daten extrahierten Wahrscheinlichkeitsmodells generiert. Da in dieser Methode kein Bezug mehr zwischen den synthetischen Daten und den echten Daten besteht, werden Herausgreifen und Verknüpfen grundsätzlich unmöglich, und Inferenz nur noch statistisch möglich. Bei sachgemäßer Verwendung kann diese Methode daher nicht nur faktische, sondern sogar absolute Anonymität erzielen, da im Gegensatz zu den anderen Anonymisierungsmethoden im synthetischen Datensatz keinerlei personenbezogene Information mehr enthalten ist, die als Angriffsfläche für einen De-Anonymisierungsversuch dienen könnte.

Für jede der Techniken wurden spezifische Kriterien entwickelt, die erfolgreiche Anonymisierung sicherzustellen. Bei der Generalisierung etwa ist dies die oben genannte K-Anonymität als ein relativ einfaches Kriterium, um Herausgreifen und Verknüpfung unmöglich zu machen, sowie L-Diversität und T-Closeness, um Inferenz zu erschweren bzw. unmöglich zu machen.

- **K-Anonymität:** Bei dieser Technik werden Quasi-Identifikatoren verschiedener Betroffener einer Vergleichseinheit mit gleichem Informationsgehalt in Gruppen zusammengefasst, sodass die dahinterstehenden Individuen nicht mehr unterscheidbar sind. Die Datensätze werden hierbei so verändert, dass jede Gruppe, die herausgegriffen werden kann, aus dem Gesamtbild mindestens eine Anzahl von k-Personen umfasst.[16] Sofern die dahinterstehenden Individuen nicht mehr getrennt werden können, ist auch eine Identifikation dieser Personen durch Verknüpfung mit korrelierendem Wissen nicht mehr eindeutig möglich. Je größer hierbei die Gruppe der zusammengefassten Informationen, desto höher ist das Maß an Anonymität und desto kleiner die Wahrscheinlichkeit als Angehöriger dieser Gruppe mit bestimmten Merkmalen identifiziert zu werden.[17]

- **L-Diversität:** Diese Technik erweitert die k-Anonymität und soll hierdurch insbesondere mögliche Angriffe mittels Inferenztechniken verhindern. Dies geschieht, in dem die einzelnen Merkmale in jeder Gruppe mindestens L verschiedene Werte aufweisen. So sollen Gruppen mit einer geringen Variabilität der Merkmalswerte begrenzt werden, damit für einen Angreifer mit Hintergrundwissen über eine bestimmte betroffene Person im Verhältnis zu dieser Gruppe bei der Identifizierung einer betroffenen Person stets eine signifikante Unsicherheit bleibt.

- **T-Closeness:** Dieser Ansatz verfeinert wiederum die L-Diversität, indem diejenigen Gruppen nachgebildet werden, die der ursprünglichen Verteilung der Merkmalswerte in der Gruppe ähneln. Hierfür wird eine weitere Bedingung für die Gruppe eingeführt: es müssen mindestens L verschiedene Werte in jeder Gruppe vertreten sein und jeder Wert muss so oft vertreten sein, dass die ursprüngliche Verteilung für jedes Merkmal abgebildet wird.

Welche konkrete Anonymisierungstechnik letztlich infrage kommt, um eine faktische Anonymisierung zu erreichen, ist dann wiederum jeweils abhängig (i) vom konkreten Datenbestand, (ii) möglicherweise verfügbarem Hintergrundwissen über die betroffene Person und (iii) dem Zweck der Auswertung[18]. Im Einzelfall sind auch Kombinationen dieser

[16] *Hansen*, in: Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht, Art. 4 Rn. 54; Artikel-29-Datenschutzgruppe, WP 216, S. 19.
[17] vgl. Stellungnahme des LDI NRW, von April 2017, Stellungnahme des LDI NRW, von April 2017, https://www.ldi.nrw.de/mainmenu_Datenschutz/submenu_Technik/Inhalt/TechnikundOrganisation/Inhalt/Anonymitaet-in-Zeiten-von-Big-Data/Anonymitaet-in-Zeiten-von-Big-Data1.pdf, zuletzt aufgerufen am 20. März 2020.
[18] Vgl. hierzu auch: Artikel-29-Datenschutzgruppe, WP 216, S. 11.

Techniken denkbar bzw. könnten notwendig sein, um das Ziel einer faktischen Anonymisierung zu erreichen.

III. Anforderungen an die Rechtsgrundlagen der Anonymisierung

Die Artikel-29-Datenschutzgruppe verlangt von Gesetzgebern, Rechtsvorschriften technisch neutral zu formulieren und „*idealerweise dem im Wandel befindlichen Entwicklungspotential der Informationstechnologie Rechnung zu tragen*“.[19] Im Hinblick auf den gesamten Bereich der Digitalisierung ist dies auch grds. zutreffend. Diese „grundsätzliche Flexibilität“ darf jedoch nicht zu einer „grundsätzlichen Rechtsunsicherheit“ führen. Dies führt stets zu einem Stillstand von Innovation und Weiterentwicklung. In Bezug auf Anonymisierungstechniken müssen den Erfindern und Anwendern von Anonymisierungstechniken daher verbindliche Leitlinien gegeben werden. Konkret müssen für alle Verfahren, die im Ergebnis ausreichend anonyme Daten nach dem faktischen Anonymisierungsbegriff hervorbringen, entsprechend einheitliche Voraussetzungen und Vorgaben definiert werden.[20] Nur so können Anwender den für sie entsprechenden Standard der Anonymisierungstechnik prüfen und wählen.[21]

IV. Rechtliche Betrachtung der Anonymisierung

Im Folgenden wird bei der rechtlichen Beurteilung jeweils von faktisch anonymisierten Daten ausgegangen. Entsprechendes gilt somit erst recht für absolut anonyme Daten, wie sie bspw. im Rahmen einer Synthetisierung entstehen.

1. Die Anonymisierung von personenbezogenen Daten als Datenverarbeitung?

In Bezug auf die Anonymisierung wird vielfach vertreten, dass diese (i) eine Datenverarbeitung darstellt und daher (ii) der DSGVO unterfällt. Dies folge daraus, dass vor der Anonymisierung personenbezogene Daten vorlagen. Jede Veränderung des Status „personenbezogenes Datum“ sei daher eine „Verarbeitung“ i.S.d. DSGVO.[22] Der Begriff sei denkbar weit auszulegen. Von der Natur der Anonymisierung heraus und unter Berücksichtigung des Schutzzwecks der DSGVO stellt eine Anonymisierung jedoch keine „Verarbeitung“ dar (siehe 1.1). Vielmehr ist die Anonymisierung eine tatsächliche Umsetzung des Grundsatzes der Datensparsamkeit und Datenminimierung (siehe 1.2) und als solche

[19] Artikel-29-Datenschutzgruppe, WP 216, S. 10.
[20] So grds. auch die Artikel-29-Datenschutzgruppe im Zusammenhang mit der Zulässigkeit der Anonymisierung als stets zulässige Weiterverarbeitung, WP 216, S. 8.
[21] Winter/Battis/Halvani: Herausforderungen für die Anonymisierung von Daten, ZD 2019, 489, 490.
[22] Vgl. hierzu: *Hansen*, in: Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht, Art. 4 Rn. 23.

keine klassische „Verarbeitung“ i.S.d. DSGVO. Auch nach der alten Rechtslage war der Bundesbeauftrage für den Datenschutz und die Informationsfreiheit („**BfDI**“) bislang der Auffassung, eine Anonymisierung stelle keine Datenverarbeitung dar (siehe 1.3.)

1.1 Die Anonymisierung ist nicht vom Schutzzweck der DSGVO umfasst

Die Anonymisierung ist eine Tätigkeit und damit auch umgangssprachlich eine „Verarbeitung“. Sie ist jedoch keine Verarbeitung im Sinne der DSGVO, da diese Tätigkeit also solche schon nicht vom Schutzzweck der DSGVO umfasst ist.

Eine Datenverarbeitung im Sinne der DSGVO ist gemäß Art. 4 Nr. 2 DSGVO

> *„Jede[r] mit oder ohne Hilfe automatisierter Verfahren ausgeführten Vorgang oder jede solche Vorgangsreihe im Zusammenhang mit personenbezogenen Daten wie das Erheben, das Erfassen, die Organisation, das Ordnen, die Speicherung, die Anpassung oder Veränderung, das Auslesen, das Abfragen, die Verwendung, die Offenlegung durch Übermittlung, Verbreitung oder eine andere Form der Bereitstellung, den Abgleich oder die Verknüpfung, die Einschränkung, das Löschen oder die Vernichtung;“*

Die dortigen Verarbeitungen haben eines gemeinsam: ihnen liegen personenbezogenen Daten zugrunde, die sowohl vor, als auch nach der Verarbeitung noch personenbezogene Daten sind.

Anonyme Daten hingegen sind keine personenbezogenen Daten (mehr). Die DSGVO ist auf anonyme Daten daher auch nicht anwendbar. Denn wo kein Personenbezug mehr vorhanden ist, muss auch keine natürliche Person geschützt werden. Die Regelungen über Datenverarbeitungen dienen ausschließlich dem Schutz des Betroffenen. Unstreitig ist die DSGVO auf Daten, die von Anfang an keinen Personenbezug aufweisen nicht anwendbar. Hier waren nie Personen betroffen, die hätten geschützt werden müssen. Dies sollte konsequenterweise auch im Hinblick auf anonymisierte Daten gelten. Im Falle einer Anonymisierung ist dem Umstand größere Bedeutung beizumessen, dass Betroffene hierbei gerade gegen eine (Re-)Identifizierung geschützt werden, als der Anwendung eines formalistischen Verarbeitungsbegriffs. Die Anonymisierung ist vielmehr die unmittelbare Umsetzung der Schutzziele der DSGVO und keine Verarbeitung.

Vor diesem Hintergrund sollte dann, wenn Daten zum Zwecke der Anonymisierung gehalten werden und dem Prozess der Anonymisierung zugeführt werden („**Anonymisierungsverarbeitung**“), nicht unter den Verarbeitungsbegriff der DSGVO fallen. Diese Daten werden zu keinem Zeitpunkt für einen anderen Zweck bereitgehalten. In diesen

Fällen würde lediglich für eine „logische juristische Sekunde“ eine Rechtsgrundlage notwendig werden und darüber hinaus einen datenschutzrechtlichen „Rattenschwanz“ nach sich ziehen, der dem Ziel der Anonymisierung und der DSGVO – dem Schutz der Betroffenen – eindeutig entgegenstehen würde.

Schließlich wäre es auch im Einklang mit der grundsätzlichen Herangehensweise der DSGVO an die Anonymisierungsverarbeitung, diese nicht als klassische Verarbeitungsvorgang einzustufen: wie unter II 1 und II 2 dargestellt, stellt die DSGVO bei ihrer Definition der Anonymisierung nicht auf die Tätigkeit der Verarbeitung von personenbezogenen Daten zu anonymen Daten (wie beim Erheben oder Löschen) ab, sondern auf das Ergebnis. Folglich sollte auch bei der rechtlichen Bewertung das Ergebnis und nicht der Weg dorthin im Fokus stehen. Dieses Ergebnis – anonymisierte Daten – ist nicht mehr im Geltungs- und Schutzbereich der DSGVO und sollte ihr daher auch insgesamt nicht unterfallen.

1.2 Die Anonymisierungsverarbeitung ist Ausprägung des Grundsatzes der Datenminimierung

Neben der Schutzfunktion erfüllt die Anonymisierungsverarbeitung insbesondere das Prinzip der Datenminimierung. Die DSGVO enthält wie schon die Richtlinie 95/46/EG insoweit sogar explizit Bestimmungen, die vorgeben, dass personenbezogene Daten standardmäßig zu anonymisieren sind.[23] Nach Art. 5 Abs. 1 lit. c DSGVO müssen personenbezogene Daten dem Zweck angemessen und erheblich sowie auf das für die Zwecke der Verarbeitung notwendige Maß beschränkt sein.

Dies gilt insbesondere, wenn der Verantwortliche personenbezogene Daten aufbewahren möchte, nachdem die ursprünglichen Verarbeitungszwecke erreicht wurden, um sie für statistische Auswertungen oder zur Forschung oder Weiterentwicklung eigener Produkte zu verwenden. In diesen Fällen sollten – sofern die Möglichkeit besteht und dadurch kein unverhältnismäßiger Aufwand entsteht – Anonymisierungstechniken zur Anwendung kommen, um eine Identifizierung unwiderruflich unmöglich zu machen.[24] Ein Beispiel einer Anonymisierungsverarbeitung ist der Einsatz synthetischer Daten für Testzwecke bei der Migration auf neue Datenverarbeitungssysteme. Im Umkehrschluss würde in diesen Fällen

[23] Vgl. zur Pseudonymisierung, was erst Recht für die Anonymisierung gelten muss: *Schwartmann/Mühlenbeck*, in: Schwartmann/Japsers/Thüsing/Kugelmann, DSGVO/BDSG, Art. 4 Rn. 66.
[24] Vgl. *Schwartmann/Mühlenbeck*, in: Schwartmann/Japsers/Thüsing/Kugelmann, DSGVO/BDSG, Art. 5 Rn. 49; *Schröder*, in: Schröder, Datenschutzrecht für die Praxis, 3. Aufl. 2019, S. 25.

eine nicht-anonymisierte Verarbeitung gegen den Grundsatz der Datenminimierung und damit die DSGVO verstoßen.[25]

Eine zu formalistische Betrachtung des Begriffs „Verarbeitung“ in Bezug auf die Anonymisierung würde diese von der DSGVO explizit gewünschte Vornahme der Anonymisierung nicht unterstützen. Sie würde zudem die grundsätzlich für Innovation und Fortschritt notwendige und der DSGVO offensichtlich auch gewollte Nutzung anonymisierter Daten aufhalten, wenn die Verantwortlichen zunächst einer umfassenden rechtlichen Evaluierung und Umsetzung gegenübersehen würden. Dies dürfte den Anreiz, Anonymisierungstechniken fortzuentwickeln und diese einzusetzen, erheblich einschränken.

1.3 Fortführung alter Rechtslage

Nach alter Rechtslage war nach Auffassung des BfDI für die Anonymisierung keine Rechtsgrundlage notwendig gewesen. Im Zusammenhang mit Big Data Projekten im Telekommunikationsbereich gab dieser an:

> *„Bei großen Datensammlungen von Telekommunikationsanbietern ist zwischen personenbezogenen und anonymisierten Datensammlungen zu unterscheiden. Eine Anonymisierung von Daten gilt nicht als Verarbeitung und ist somit zulässig.“*[26]

Dieser Bericht erschien zum Bundesdatenschutzgesetz („**BDSG**“) aF sowie zur Datenschutz-Richtlinie 95/46. Der Verarbeitungsbegriff wurde in der DSGVO jedoch nahezu wortgleich übernommen. Auch galten nach neuem Recht dieselben Faktoren zur Bestimmung, ob eine Anonymisierung vorlag, wie schon nach dem BDSG aF und der Datenschutzrichtlinie 95/46/EG.[27]

Eine Änderung der Rechtslage, aus der sich nunmehr ein Erfordernis einer Rechtsgrundlage ergeben sollte, ist daher nicht erkennbar.

2. Privilegierung wirksamer Anonymisierungstechniken

Sofern entgegen des Schutzzieles der DSGVO und Vorgaben zur Datenminimierung dennoch von einer Datenverarbeitung ausgegangen wird, ist hinsichtlich der

[25] *Schwartmann/Mühlenbeck*, in: Schwartmann/Japsers/Thüsing/Kugelmann, DSGVO/BDSG, Art. 5 Rn. 49.
[26] 26. Tätigkeitsbericht zum Datenschutz für die Jahre 2015/2016, 17.2.4.4 Big Data im TK-Bereich, S: 171.
[27] Vgl. hierzu *Ernst*, in: Paal/Pauly, DS-GVO BDSG, 2. Aufl. 2018, Art. 4 Rn. 450.

Anonymisierungsverarbeitung eine Privilegierung notwendig. Nur so lassen sich die zuvor dargestellten Schutzziele der DSGVO und der Grundsatz der Datenminimierung erreichen, bzw. zur Erreichung dieser Ziele durch Entwicklung entsprechender Anonymisierungstechniken entsprechende Anreize schaffen.

Eine solche Privilegierung würde zudem langfristig in der Praxis dazu führen, dass personenbezogene Daten nicht auf Basis möglicherweise konstruierter Rechtsgrundlagen oder Einwilligungen „weiterverarbeitet" werden. Einwilligungen, die häufig trotz umfangreicher Erklärungen leichtfertig erteilt werden könnten – was wiederum nicht im Interesse der Betroffen oder der DSGVO ist. Eine privilegierte anonyme „Weiterverarbeitung" wäre folglich stets im Interesse aller Beteiligten.

Eine solche Privilegierung sollte mindestens das Erfordernis einer Rechtsgrundlage entfallen lassen und einfachere Anforderungen an die Erfüllung der Informationspflichten stellen. So sollte der Verantwortliche lediglich verpflichtet sein, darauf hinzuweisen, dass Anonymisierungstechniken angewendet werden, ohne auf diese Techniken oder die weitergehende Nutzung der anonymen Daten eingehen zu müssen.

3. Anonymisierung ist stets datenschutzrechtlich zulässig

Unter der Prämisse, dass Daten im Einzelfall durch eine Anonymisierungstechnik faktisch anonymisiert werden, steht am Ende immer dasselbe Ergebnis: Anonyme Daten (siehe 3.1). Folglich ist Anonymisierungsverarbeitung – sofern sie eine „Verarbeitung" darstellt – auch immer nach denselben rechtlichen Kriterien zu beurteilen und daher immer datenschutzrechtlich zulässig (siehe 3.2 und 3.3).

3.1 Anonymisierung ist kein Einzelfall

Rechtlich ist die Anonymisierungsverarbeitung im Allgemeinen zu betrachten und auf dieses Ergebnis (nämlich die Erstellung anonymer Daten) zu beschränken. Anonymisierungsverarbeitungen erzeugen immer dasselbe Ergebnis. Die anschließenden Zwecke der Verwendung dieser Daten können hierbei keine Rolle mehr spielen. Auch die Artikel-29-Datenschutzgruppe geht insoweit davon aus, dass die Anonymisierung als eine Form der Weiterverarbeitung personenbezogener Daten stets mit dem ursprünglichen

Verarbeitungszweck vereinbar ist, sofern das Anonymisierungsverfahren geeignet ist, zuverlässig anonymisierte Informationen hervorzubringen.[28]

Sofern die Artikel-29-Datenschutzgruppe z.B. davon spricht, dass das Anonymisierungsverfahren im Hinblick auf seinen Zweck geprüft werden muss (siehe oben unter II 1), ist damit gemeint, ob das Verfahren unter Berücksichtigung dieses Zwecks, z.B. Veröffentlichung der anonymen Daten, tatsächlich geeignet ist, anonyme Daten herzustellen. Die Einzelfallprüfung bezieht sich also jeweils auf die Auswahl der Anonymisierungstechnik – nicht auf die rechtliche Grundlage. So können z.B. Daten im nichtöffentlichen Raum anonym sein, veröffentlicht für jemandem mit dem notwendigen Zusatzwissen, jedoch wiederum ggf. bestimmbar. Diese Prüfung ist jedoch unabhängig davon, ob eine grundsätzlich – auf der ersten faktischen Stufe – erfolgreiche Anonymisierungsverarbeitung auch rechtlich zulässig ist.

3.2 Art. 6 Abs. 4 DSGVO i.V.m. der ursprünglichen Rechtsgrundlage

Eine Anonymisierungsverarbeitung ist folglich immer nach Art. 6 Abs. 4 DSGVO zulässig, sofern es sich um die Anonymisierung bereits vorhandener personenbezogener Daten, die zu anderen Zwecken erhoben wurden, handelt. Wie zuvor dargestellt, ist dies auch die Auffassung der Artikel-29-Datenschutzgruppe. Art. 6 Abs. 4 DSGVO lautet:

> *„Beruht die Verarbeitung zu einem anderen Zweck als zu demjenigen, zu dem die personenbezogenen Daten erhoben wurden, nicht auf der Einwilligung der betroffenen Person oder auf einer Rechtsvorschrift der Union oder der Mitgliedstaaten, die in einer demokratischen Gesellschaft eine notwendige und verhältnismäßige Maßnahme zum Schutz der in Artikel 23 Absatz 1 genannten Ziele darstellt, so berücksichtigt der Verantwortliche – um festzustellen, ob die Verarbeitung zu einem anderen Zweck mit demjenigen, zu dem die personenbezogenen Daten ursprünglich erhoben wurden, vereinbar ist – unter anderem*
>
> *a) jede Verbindung zwischen den Zwecken, für die die personenbezogenen Daten erhoben wurden, und den Zwecken der beabsichtigten Weiterverarbeitung,*
>
> *b) den Zusammenhang, in dem die personenbezogenen Daten erhoben wurden, insbesondere hinsichtlich des Verhältnisses zwischen den betroffenen Personen und dem Verantwortlichen,*
>
> *c) die Art der personenbezogenen Daten, insbesondere ob besondere Kategorien personenbezogener Daten gemäß Artikel 9 verarbeitet werden*

[28] Artikel-29-Datenschutzgruppe, WP 216, S. 8.

oder ob personenbezogene Daten über strafrechtliche Verurteilungen und Straftaten gemäß Artikel 10 verarbeitet werden,

d) die möglichen Folgen der beabsichtigten Weiterverarbeitung für die betroffenen Personen,

d) das Vorhandensein geeigneter Garantien, wozu Verschlüsselung oder Pseudonymisierung gehören kann."

Der Kern der Vorschrift ist die Prüfung der Vereinbarkeit von ursprünglichem und neuem Zweck unter Berücksichtigung der hierin (nicht abschließend) genannten Kriterien.[29] Folglich müssen für die Beurteilung einer kompatiblem Zweckänderung nicht alle Kriterien erfüllt sein.[30]

Im Einzelnen:

- **Verbindung zwischen ursprünglichem und neuem Zweck (Anonymisierung)**: Ergibt sich danach, dass die Weiterverarbeitung ein logischer nächster Schritt in Bezug auf den ursprünglichen Zweck ist, spricht dies für eine Vereinbarkeit der Zwecke.[31]

- **Zusammenhang bei der ursprünglichen Erhebung**: Hierbei ist insbesondere das Verhältnis zwischen den betroffenen Personen und dem Verantwortlichen maßgeblich. Dabei ist vor allem auf die vernünftigen Erwartungen („reasonable expectations") der betroffenen Person abzustellen, die sich hinsichtlich der weiteren Verwendung ihrer Daten aus der Beziehung zu dem Verantwortlichen ergeben.[32]

- **Art der personenbezogenen Daten**: Dies ist kein wirkliches Kriterium der Prüfung, sondern unterstreicht, dass im Falle der Weiterverarbeitung besonderer Kategorien personenbezogener Daten, eine entsprechend sorgfältige Prüfung erfolgen sollte.[33]

- **Folgen der Weiterverarbeitung**: Hierbei ist zu berücksichtigen, welche Risiken mit der beabsichtigten Verarbeitung für den anderen Zweck verbunden sind, sowie

[29] *Schwartmann/Mühlenbeck*, in: Schwartmann/Japsers/Thüsing/Kugelmann, DSGVO/BDSG, Art. 5 Rn. 188.
[30] *Schwartmann/Mühlenbeck*, in: Schwartmann/Japsers/Thüsing/Kugelmann, DSGVO/BDSG, Art. 5 Rn. 190.
[31] *Heberlein*, in: Ehmann/Selmayr, Datenschutz-Grundverordnung, 2. Aufl. 2018, Art. 6 Rn. 55.
[32] *Heberlein*, in: Ehmann/Selmayr, Datenschutz-Grundverordnung, 2. Aufl. 2018, Art. 6 Rn. 56.
[33] *Heberlein*, in: Ehmann/Selmayr, Datenschutz-Grundverordnung, 2. Aufl. 2018, Art. 6 Rn. 58.

deren Eintrittswahrscheinlichkeit und Schwere für die Rechte und Freiheiten der betroffenen Person.[34]

- **Vorhandensein geeigneter Garantien**: Besondere Bedeutung kommt den Schutzmaßnahmen zu, die die Zweckänderung der Sache, nach mit den Interessen und dem Schutz der betroffenen Person in Einklang bringen. Dies erfordert, dass der Verantwortliche dem mit der Verarbeitung verbundenen Risiko angemessene organisatorische und technische Maßnahmen trifft, die in der konkreten Verarbeitungssituation den effektiven Schutz der betroffenen Person gewährleisten.

Diese Kompatibilitätsprüfung des Art. 6 Abs. 4 DSGVO geht davon aus, dass personenbezogene Daten weiterverarbeitet und im Anschluss für einen neuen Zweck genutzt werden. Daher die Prüfung eng am ursprünglichen Zweck der Verarbeitung, damit die anschließende Nutzung der Daten nicht unverhältnismäßig ausgeweitet wird und der Betroffene dadurch nicht die Kontrolle über seine Daten verliert. So kommt gerade eine Verarbeitung dann in Betracht, je stärker der Personenbezug der Daten verringert wird.[35] Dies muss bei der Kompatibilitätsprüfung hinsichtlich der Weiterverarbeitung in Form der Anonymisierung entsprechend berücksichtig werden, da hier im Ergebnis gerade keine personenbezogenen Daten mehr verarbeitet werden. Dies hat (positive) Auswirkungen insbesondere auf die Prüfung der möglichen Folgen und dem Vorhandensein geeigneter Garantien.

- Die Folgen einer Anonymisierungsverarbeitung sind für den Betroffenen denkbar gering. Im Falle einer absoluten Anonymisierung (wie bspw. bei der Erzeugung synthetischer Daten) entstehen keinerlei Folgen für den Betroffenen. Die Verbindung zu seinen Daten wird vollständig entfernt. Aber auch bei einer (ausreichenden) faktischen Anonymisierung sind die Folgen für Betroffenen äußerst gering. Hierbei ist nur in seltenen Fällen unter Einsatz entsprechend unverhältnismäßiger Mittel ggf. eine De-Anonymisierung möglich.

- Mit der Wahl der jeweils geeigneten Anonymisierungstechnik durch den Verantwortlichen wird auch garantiert, dass eine Methode gewählt wird, die eine De-Anonymisierung so weit wie möglich ausschließt. Dies ist insbesondere ausreichend,

[34] *Heberlein*, in: Ehmann/Selmayr, Datenschutz-Grundverordnung, 2. Aufl. 2018, Art. 6 Rn. 59.
[35] *Frenzel*, in: Paal/Pauly, DS-GVO BDSG, 2. Aufl. 2018, Art. 6 Rn. 52.

da nach dem Wortlaut bereits Pseudonymisierungstechniken oder Verschlüsselungstechniken solche Garantien sein können.

- Vor dem Hintergrund, dass die DSGVO beim Prinzip der Datenminimierung bereits statuiert, dass sofern und soweit möglich anonyme oder pseudonyme Daten verwendet werden sollen, stellt eine Anonymisierungsverarbeitung einmal erhobener Daten auch eine „logische Folge“ der ursprünglichen Verarbeitung dar.

Schließlich sind die in Art. 6 Abs. 4 DSGVO genannten Kriterien nach dem Willen des Gesetzgebers nicht im Kontext jeder zweckändernden Verarbeitung zu berücksichtigen. Weiterverarbeitungen für im öffentlichen Interesse liegende Archivzwecke, für wissenschaftliche und historische Forschungszwecke oder statistische Zwecke, sind insoweit in Art. 5 Abs. 1 lit. b Hs. 2 DSGVO privilegiert und fallen demnach aus dem normativen Anwendungsbereich des stets erforderlichen Kompatibilitätstests heraus, sodass auch die in Art. 6 Abs. 4 festgelegten Kriterien in diesen Fällen nicht greifen können.[36] Art. 5 Abs. 1 lit. b Hs. 2 DSGVO verweist wiederum auf Art. 89 Abs. 1 DSGVO. Dieser Test privilegiert Verarbeitungen, die geeignete Garantien für die Rechte und Freiheiten der betroffenen Person gemäß der DSGVO unterliegen. Mit diesen Garantien werde sichergestellt, dass technische und organisatorische Maßnahmen bestehen, mit denen insbesondere die Achtung des Grundsatzes der Datenminimierung gewährleistet werde. Zu diesen Maßnahmen kann die Pseudonymisierung gehören, sofern es möglich ist, diese Zwecke auf diese Weise zu erfüllen. Wenn schon in diesen Fällen eine Verarbeitung pseudonymisierter Daten (also nach wie vor personenbezogener Daten) privilegiert und von der Kompatibilitätsprüfung in Art. 6 Abs. 4 DSGVO ausgenommen ist, dann muss eine entsprechende Privilegierung erst recht für die Anonymisierungsverarbeitung gelten.

Entgegen der Vorgehensweise im Beispiel des BfDI ist jeweils der ursprüngliche Erhebungszweck der Daten mit dem Weiterverarbeitungszweck der Anonymisierung ins Verhältnis zu setzen. Im Beispiel des Positionspapieres des BfDI wurden jedoch (i) die Anonymisierung der Daten und (ii) die letztlich bezweckte Optimierung der Dienstleistungen (anhand der anonymisierten Daten) mit dem ursprünglichen Zweck der Erhebung (Vertragserfüllung) in Beziehung gesetzt. Folglich wird auch die weitere Verarbeitung der anonymisierten Daten in die Zweckänderungsprüfung einbezogen. Dies ist jedoch dogmatisch falsch. Denn – wie festgestellt – unterfällt eine (weitere) Verarbeitung

[36] *Albers/Veit*, in: BeckOK Datenschutzrecht, Wolff/Brink, 30. Edition, Stand: 01.11.2019, Art. 6 Rn. 70.

anonymisierter Daten nicht (mehr) der DSGVO. Daher darf sie auch nicht Bestandteil einer Verhältnismäßigkeitsprüfung der DSGVO sein. Die Situation ist vielmehr wie folgt.

- **Schritt 1:** ursprüngliche Verarbeitung (inklusive Rechtsgrundlage)
- **Schritt 2:** Anonymisierung als Weiterverarbeitung (Rechtsgrundlage Art. 6 Abs. 4 DSGVO)
- **Schritt 3:** darauffolgende „Verarbeitung" der anonymisierten Daten für andere Zwecke (keine Rechtsgrundlage mehr notwendig nach DSGVO)

Die Frage, inwiefern es für die Anonymisierungsverarbeitung einer Rechtsgrundlage bedarf, darf sich also auch nur auf den eigentlichen Akt der Anonymisierung beziehen. Jede darauffolgende Nutzung der anonymisierten Daten ist ohne weitere datenschutzrechtliche Grundlage möglich.

Die Anonymisierungsverarbeitung ist – wie dargelegt – grundsätzlich nach Art. 6 Abs. 4 mit dem Zweck der ursprünglichen Datenverarbeitung kompatibel.

3.3 Art. 6 Abs. 1 lit. f DSGVO: berechtigte Interessen des Verantwortlichen

Auch Art. 6 Abs. 1 lit. f DSGVO kann regelmäßig Rechtsgrundlage für die Anonymisierungsverarbeitung von Daten sein. Dieser Auffassung ist auch die Artikel-29-Datenschutzgruppe in ihrem Working Paper 216. Dies ist insbesondere dann der Fall, wenn die zu anonymisierenden personenbezogenen Daten nicht für einen anderen vorherigen Zweck bereits erhoben waren, sondern ausschließlich zum Zweck der Anonymisierungsverarbeitung erhoben werden oder ganz neu verwendet werden. Beispielsweise kann sich ein Auftragsverarbeiter, der die Daten seines Verantwortlichen Auftraggebers für die Verwendung eigener Zwecke anonymisiert (um seine Künstliche Intelligenz weiterzuentwickeln), nicht auf Art. 6 Abs. 4 DSGVO berufen.

Gemäß Art. 6 Abs. 1 lit. f DSGVO gilt zur Rechtmäßigkeit der Verarbeitung:

> *„die Verarbeitung ist zur Wahrung der berechtigten Interessen des Verantwortlichen oder eines Dritten erforderlich, sofern nicht die Interessen oder Grundrechte und Grundfreiheiten der betroffenen Person, die den Schutz personenbezogener Daten erfordern, überwiegen, insbesondere dann, wenn es sich bei der betroffenen Person um ein Kind handelt."*

Die Vornahme einer Anonymisierung zur weiteren freien Nutzung der Daten ohne Personenbezug ist stets ein legitimes Interesse des Verantwortlichen. Sie ist hierfür – wie dargestellt – sogar nach der DSGVO erforderlich. In Bezug auf die hier geforderte Interessenabwägung zwischen dem berechtigten Interesse des Verantwortlichen und den Rechten und Freiheiten der Betroffenen kann nichts Anderes als das zuvor zur Kompatibilitätsprüfung Gesagte gelten. Insbesondere das Risiko für die Betroffenen bei der Verwendung der Daten im Anschluss an die Anonymisierungsverarbeitung ist minimal. Berücksichtigt werden muss bei dieser Interessenabwägung insbesondere wieder, dass im Anschluss gerade keine personenbezogenen Daten mehr genutzt werden. Die Anonymisierung erfolgt auch zum Schutz des Betroffenen, dessen Rechte und Freiheiten in diesem Falls daher nicht überwiegen.

4. Informationspflichten bei der Anonymisierung

Bislang noch nicht im Positionspapier des BfDI thematisiert wurden der Umgang mit den Informationspflichten, die zwangsläufig folgen, wenn eine Datenverarbeitung nach der DSGVO vorliegt.

Eine Verarbeitung personenbezogener Daten für einen anderen Zweck setzt, neben der Bedingung eines Rechtsgrundes für die Verarbeitung zu den anderen Zwecken auch die Einhaltung der anderen Grundsätze der DSGVO voraus. Dies umfasst insbesondere die Information der betroffenen Person über diese neuen Zwecke und ihre Rechte.[37] Dementsprechend ist der Verantwortliche gemäß Art. 13 Abs. 3 DSGVO und Art. 14 Abs. 4 DSGVO verpflichtet, die betroffene Person vor der von ihm beabsichtigten Weiterverarbeitung für einen anderen Zweck auch über diesen anderen Zweck zu informieren und ihr alle anderen maßgeblichen Informationen zu geben, die für die Gewährleistung einer fairen und transparenten Verarbeitung notwendig sind.

Wie zuvor bereits angeregt, kommt auch hier nur eine Privilegierung dieser Pflichten in Betracht. Es sollte ausreichend sein, wenn der Verantwortliche, die durch die Anonymisierung Betroffenen, durch einen allgemeinen Hinweis darüber informiert, dass er die personenbezogenen Daten ggf. (i) einem Anonymisierungsverfahren unterziehen wird und (ii) diese im Anschluss nicht mehr auf eine natürliche Person zurückzuführen sind.

[37] *Heberlein*, in: Ehmann/Selmayr, Datenschutz-Grundverordnung, 2. Aufl. 2018, Art. 6 Rn. 54.

Weitergehende Informationen hinsichtlich der anschließenden Verwendung der anonymen Daten sind hingegen nicht mehr erforderlich, da diese nicht mehr der DSGVO unterfallen.

* * *