FACHBEREICH
SICHERHEIT – SCHUTZ
UND ZUVERLÄSSIGKEIT

Stellungnahme: Anonymisierung unter der DSGVO

Stellungnahme des Fachbereiches Sicherheit der Gesellschaft für Informatik e.V. (GI) zum Konsultationsverfahren des Bundesbeauftragten für den Datenschutz und die Informationsfreiheit (BfDI)

Stand: 08.03.2020

Autoren:

- Prof. Dr. Ralf Kneuper, IUBH Internationale Hochschule, r.kneuper@iubh-fernstudium.de (Ansprechpartner)
- Marion Steiner, IT-Security@Work GmbH, marion.steiner@isw-online.de
- Benedict Voßbein, UIMC Dr. Voßbein GmbH & CO. KG

Reviewer:

- Georg Reiss, Stadtwerke Frankfurt am Main Holding GmbH, g.reiss@stadtwerke-frankfurt.de
- Bernhard C. Witt, it.sec GmbH, Sprecher des GI-Fachbereichs Sicherheit, bcwitt@it-sec.de

Hintergrund

Im Rahmen eines öffentlichen Konsultationsverfahrens hat der BfDI ein Dokument zur „Anonymisierung unter der DSGVO unter besonderer Berücksichtigung der TK-Branche“ veröffentlicht und zu Kommentaren und Stellungnahmen dazu aufgefordert.[1]

Das vorliegende Dokument gibt die Stellungnahme des Fachbereiches *Sicherheit – Schutz und Zuverlässigkeit* der Gesellschaft für Informatik e.V. zu diesem BfDI-Dokument wieder.

Zusammenfassung

Das vorliegende Dokument des BfDI kommt im Wesentlichen zu dem Ergebnis, dass eine Anonymisierung eine geeignete Rechtsgrundlage erfordert, und nennt als mögliche solche Rechtsgrundlagen die Einwilligung (Art. 6 Abs. 1 lit. a DSGVO), die Weiterverarbeitung (Art. 6 Abs. 4 DSGVO i.V.m. der ursprünglichen Rechtsgrundlage) sowie die Erfüllung rechtlicher Pflichten (Art. 6 Abs. 1 Buchst. c) i.V.m. Art. 17 Abs. 1 DSGVO).

Hier kommt nach Einschätzung der Autoren dieser Stellungnahme zusätzlich das berechtigte Interesse als Rechtsgrundlage in Frage.

Eine wesentliche Einschränkung bei allen diesen Rechtsgrundlagen ist aber, dass sie einen ausreichenden Grad der Anonymisierung erfordern, der im aktuellen Dokument des BfDI unter dem Stichwort der *Validität* zwar angesprochen, aber nicht tiefer behandelt ist.

Dabei ist eine Unterscheidung zweier verschiedener Rollen der Anonymisierung erforderlich. Einerseits handelt es sich um eine Verarbeitung von Daten, die wie beschrieben eine Rechtsgrundlage benötigt. Andererseits handelt es sich um eine Maßnahme zum Datenschutz, wobei der Grad der erreichten oder zu erreichenden Anonymisierung betrachtet werden muss, denn es handelt sich bei der Anonymisierung nicht um eine einfache Ja-Nein-Eigenschaft, die vorhanden ist oder auch nicht. Es stellt sich also jeweils die Frage nach dem im Einzelfall gegebenen Grad der Anonymisierung, der auch wieder deutlichen Einfluss darauf hat, ob eine vorgesehene Rechtsgrundlage tatsächlich gegeben ist.

Die Frage nach dem angemessenen oder erforderlichen Grad der Anonymisierung sollte daher in einer Überarbeitung des BfDI-Dokuments ausführlicher berücksichtigt werden. Diese Betrachtung muss ausgehen von den mit der Verarbeitung der anonymisierten Daten verbleibenden Risiken, auch im Hinblick auf die Art der zu verarbeitenden Daten (beispielsweise besondere Kategorien personenbezogener Daten), und mit Hilfe von verbreiteten Anonymitätsmaßen wie k-Anonymität und ε-Differential Privacy die Anforderungen konkretisieren.

[1] `https://www.bfdi.bund.de/SiteGlobals/Modules/Buehne/DE/Startseite/Termin_Link/HP_Text_Termin.html`

1 Einführung

Ein wesentlicher Aspekt bei einer Betrachtung von Anonymisierung unter der DSGVO ist eine Unterscheidung zwischen zwei Sichtweisen der Anonymisierung: Anonymisierung ist einerseits eine Form der Verarbeitung personenbezogener Daten, für die sich die Frage nach der Rechtsgrundlage und der Einhaltung anderer Rahmenbedingungen stellt. Daneben ist Anonymisierung aber eine Maßnahme zur Umsetzung des Datenschutzes nach DSGVO, bei der sich Fragen danach stellen, ob bzw. wann diese Maßnahme angemessen und ausreichend ist. Im Folgenden werden diese beiden Sichtweisen separat betrachtet, wobei aber zu berücksichtigen ist, dass diese gleichzeitig gelten, d.h. Anonymisierung ist nicht entweder Verarbeitung oder Maßnahme, sondern im Normalfall beides gleichzeitig.

Anonymisierung als Maßnahme ist in der DSGVO nicht näher beschrieben. Sie liegt aber in ihrem Schutzumfang zwischen Pseudonymisierung und Löschung der Daten, so dass die Aussagen über diese Maßnahmen einen Rahmen für die Bewertung darstellen. Anonymisierung ist einerseits vergleichbar mit der Pseudonymisierung. Bei beiden Maßnahmen (sofern sie erfolgreich durchgeführt werden) haben Dritte keine Möglichkeit der Zuordnung der Daten zu einer Person. Bei einer Pseudonymisierung hat allerdings der Verantwortliche noch die Möglichkeit der Zuordnung, während es aus Sicht von Dritten keinen Unterschied zwischen beiden Maßnahmen gibt.

Anonymisierung ist andererseits vergleichbar mit der Löschung von Daten, denn beide führen zu dem Ergebnis, dass die verarbeiteten Daten anschließend nicht mehr als personenbezogene Daten zur Verfügung stehen. Allerdings ist das Risiko einer Re-Identifikation anonymisierter Daten höher als das Risiko einer Wiederherstellung gelöschter Daten, so dass beispielsweise Roßnagel bezweifelt, dass eine Löschungsverpflichtung durch eine Anonymisierung erfüllt werden kann.[2]

2 Anonymisierung als Verarbeitung personenbezogener Daten

Bei der Betrachtung von Anonymisierung als Verarbeitung von Daten stellen sich die beiden vom BfDI genannten Fragen:

- Handelt es sich bei der Anonymisierung um eine Verarbeitung von Daten, für die eine Rechtsgrundlage erforderlich ist?
- Welche Rechtsgrundlage kommt dafür ggf. dafür in Frage?

Gemäß dem BfDI-Dokument ist für Anonymisierung eine Rechtsgrundlage erforderlich; da es sich um eine Weiterverarbeitung ohne Zweckänderung handelt, kann die Anonymisierung in vielen Fällen auf Basis der Rechtsgrundlage der ursprünglichen Verarbeitung geschehen.

[2] Simitis/Hornung/Spiecker gen. Döhmann (2019): *Datenschutzrecht*. Nomos Verlag. DSGVO Art. 4 Nr. 2 Begriffsbestimmung „Verarbeitung“ Randnummer 32

2.1 Notwendigkeit einer Rechtsgrundlage

Die Einschätzung des BfDI scheint angemessen. Der Begriff der Verarbeitung gemäß DSGVO ist weit zu fassen, und bei der Anonymisierung handelt es sich um eine ähnliche Form der Verarbeitung wie die Löschung, die in Art. 4 Abs. 2 DSGVO explizit als Verarbeitung genannt ist, auch wenn die Anonymisierung dort nicht genannt ist. Damit greift Art. 6 DSGVO, und Anonymisierung erfordert eine Rechtsgrundlage.

2.2 Identifikation der Rechtsgrundlage

Auch hier erscheinen die vom BfDI genannten möglichen Rechtsgrundlagen der Einwilligung (Art. 6 Abs. 1 lit. a DSGVO), der Weiterverarbeitung (Art. 6 Abs. 4 DSGVO i.V.m. der ursprünglichen Rechtsgrundlage) sowie der Erfüllung rechtlicher Pflichten (Art. 6 Abs. 1 lit. c i.V.m. Art. 17 Abs. 1 DSGVO) grundsätzlich angemessen.

Zusätzlich zu den genannten Rechtsgrundlagen kommt noch das berechtigte Interesse (Art. 6 Abs. 1 lit. f DSGVO) in Frage. In diesem Fall sind die üblichen Kriterien zu prüfen, insbesondere die Abwägung der Kriterien der Betroffenen gegen die des Verantwortlichen. Mit wenigen Ausnahmen wird es aber kein besonderes Interesse der Betroffenen geben, das gegen eine Anonymisierung spricht, so dass eine Anonymisierung auf dieser Rechtsgrundlage meist möglich sein wird.

Dabei ist allerdings der unterschiedliche Charakter dieser Rechtsgrundlagen zu beachten: Während bei den meisten dieser Rechtsgrundlagen die Anonymisierung in erster Linie den Charakter einer Maßnahme zum Schutz der Betroffenen hat, hat sie bei der Erfüllung rechtlicher Pflichten den Charakter, die Pflicht zur Löschung zu ersetzen. Dieser andere Charakter führt dazu, dass Roßnagel wie oben beschrieben zum Schluss kommt, dass diese Rechtsgrundlage nicht trägt, da Löschen nach seiner Einschätzung wegen des höheren damit verbundenen Risikos nicht durch Anonymisierung erfolgen kann.

Voraussetzung für die Nutzung dieser Rechtsgrundlagen ist, dass es sich um eine *effektive* Anonymisierung handelt, die nicht mit vertretbarem Aufwand rückgängig zu machen ist. Hier kommt die Frage nach der Anonymisierung als Maßnahme zum Datenschutz ins Spiel, wie im folgenden Kapitel erläutert. Dabei sei darauf hingewiesen, dass auch eine Löschung in vielen Fällen nicht absolut ist, sondern mit entsprechendem Aufwand wieder aufgehoben werden kann, beispielsweise durch Zugriff auf verbleibende Sicherungskopien. oder Archivversionen. Das Risiko einer unzureichenden Anonymisierung ist allerdings wesentlich größer als das einer unzureichenden Löschung, da in diesem Fall wesentliche Teile der Daten in diesem Fall bewusst weiterhin existieren und bereitgestellt werden.

Informations- und Auskunftsrechte der Betroffenen

Bei den meisten der identifizierten Rechtsgrundlagen ist zu berücksichtigen, dass hier gemäß Art. 13 bzw. Art. 14 DSGVO eine Information der Betroffenen über die geplante Verarbeitung, also die Anonymisierung, erforderlich ist. Es ist daher wichtig, auf die geplante Anonymisierung schon von Beginn an hinzuweisen. Eine Ausnahme ist die Anonymisierung auf Basis von Garantien und Ausnahmen zu statistischen Zwecken (Art. 89),

wenn die ursprünglichen personenbezogenen Daten nicht beim Betroffenen erhoben wurden (Art. 14 Abs. 5 lit. b).

Ein Auskunftsrecht der Betroffenen über die anonymisierten Daten besteht dagegen gemäß Art. 11 DSGVO nicht, da die Zuordnung der Daten zum Betroffenen durch die Anonymisierung gerade nicht mehr möglich ist.

3 Anonymisierung als Maßnahme zum Datenschutz

Neben der oben beschriebenen Rolle der Anonymisierung als Verarbeitung von Daten handelt, die möglicherweise eine Rechtsgrundlage erfordert, hat Anonymisierung auch die Rolle einer Maßnahme zum Datenschutz. Dieser Aspekt ist im Papier des BfDI nur angerissen, spielt aber eine wesentliche Rolle, die auch wieder auf die Rechtmäßigkeit der Anonymisierung rückwirkt. Auch hier gibt es zwei grundlegende Fragen:

- Wann ist Anonymisierung als Maßnahme zum Datenschutz angemessen?
- Wann ist Anonymisierung als Maßnahme zum Datenschutz ausreichend?

Im Folgenden werden diese Fragen genauer betrachtet, und sie sollten auch in einer Überarbeitung des BfDI-Dokumentes berücksichtigt werden.

3.1 Wann ist Anonymisierung als Maßnahme zum Datenschutz angemessen?

Angemessene Gründe für die Nutzung von Anonymisierung wurden bereits bei der Frage der Rechtmäßigkeit angesprochen. Wichtigster Grund ist, personenbezogene Daten statistisch oder in ähnlicher Form auszuwerten, wobei der Personenbezug für diese Auswertungen nicht erforderlich ist, oder evtl. gegen die DSGVO verstoßen würde.

Grundsätzlich kann man davon ausgehen, dass eine Anonymisierung überall dort in Frage kommt, wo gemäß DSGVO eine Pseudonymisierung angemessen ist. Wie in der Einleitung beschrieben kann Anonymisierung als verstärkte Form der Pseudonymisierung betrachtet werden. Umgekehrt bedeutet das, dass die im Folgenden beschriebenen Betrachtungen zum erforderlichen Grad der Anonymisierung im Wesentlichen auch für Pseudonymisierung gelten, da für Beteiligte, die keinen Zugang zu den Zuordnungsinformationen haben, kein Unterschied zwischen anonymen und pseudonymen Daten bestehen soll.

Schwieriger ist es bei der Anonymisierung als Ersatz für eine Löschung, da bei einer Anonymisierung ein Teil der Daten weiterhin vorhanden ist und dadurch das Risiko einer Re-Identifikation größer ist als das Risiko einer Wiederherstellung gelöschter Daten. Roßnagel kommt daher zum Schluss, dass Löschen nicht durch Anonymisierung erfolgen kann. Selbst wenn man dieser Meinung nicht folgt, dann ist doch eindeutig, dass eine Anonymisierung als Ersatz für Löschung entsprechend hohe Anforderungen an die Qualität der Anonymisierung stellt, die im BfDI-Dokument bisher nicht ausreichend berücksichtigt sind und im folgenden Abschnitt betrachtet werden sollen.

Soweit aber nicht explizit die Löschung der Daten gefordert ist, erscheint unstrittig, dass Anonymisierung als Maßnahme in Frage kommt, beispielsweise zur Umsetzung von Datenschutz-Prinzipien wie Datenminimierung und Speicherbegrenzung.

3.2 Wann ist Anonymisierung als Maßnahme zum Datenschutz ausreichend?

Wie im BfDI-Papier auf Basis von ErwG 26 DSGVO beschrieben, wird keine absolute Anonymität gefordert, aber die Re-Identifizierung muss angemessen schwierig sein, genauer gesagt sollen bei der Bewertung „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“ (ErwG 26 DSGVO).

Das führt u.a. zu der Frage, welche „anderen Personen“ hier betrachtet werden sollen. Dabei muss man nach allgemeinem Ermessen sicher davon ausgehen, dass Unternehmen, bei denen Datenanalyse ein wesentlicher Baustein des Geschäftsmodells ist und bei denen damit entsprechend großes Zusatzwissen über Betroffene bereits vorhanden ist, wahrscheinlich sehr viel weitreichendere Mittel nutzen können und werden als Unternehmen mit einem anderen Geschäftsmodell, sofern sie Zugriff auf die anonymisierten Daten bekommen.

Damit kommen wir zu der zentralen Feststellung, dass Anonymität keine Ja/Nein-Eigenschaft ist und es daher „die“ Anonymisierung bzw. Anonymität nicht gibt, sondern eine Betrachtung der unterschiedlichen Grade der erreichten bzw. der geforderten Anonymität erforderlich ist. Diese verschiedenen Grade spiegeln sich beispielsweise in den verschiedenen Anonymitätsmodellen und den dort verwendeten Parametern wieder, siehe beispielsweise k-Anonymität sowie deren Erweiterungen (z.B. i-Diversität) oder ε-Differential Privacy.

Weiterhin folgt daraus, dass ggf. trotz Anonymisierung ein Schutz der (ehemals personenbezogenen) Daten weiterhin erforderlich ist, wenn eine Re-Identifizierung mit den Mitteln eines Datenanalyse-Unternehmens möglich wäre. Es gibt eine Reihe öffentlich bekannter Beispiele, wo Daten anonymisiert und dann veröffentlicht wurden, diese Daten dann aber mit überschaubarem Aufwand zumindest teilweise wieder re-identifiziert werden konnten

Abhängig von der verwendeten Rechtsgrundlage gibt es nun verschiedene Gründe, warum der Grad der Anonymisierung, ausgehend von der beabsichtigten Verarbeitung der anonymisierten Daten, bewertet werden und ausreichend hoch sein muss:

- Handelt es sich um eine Verarbeitung auf der Rechtsgrundlage einer Einwilligung, so muss diese in informierter Weise abgegeben werden (Art. 4 Abs. 11 DSGVO). Mindestens wenn diese Information ein ggf. existierendes hohes Risiko einer Re-Identifizierung nicht berücksichtigt, dann ist die Gültigkeit der Einwilligung in Frage zu stellen.

- Handelt es sich um eine Weiterverarbeitung, so fordert Art. 6 Abs. 4 lit. d eine Berücksichtigung der „möglichen Folgen der beabsichtigten Weiterverarbeitung für die

betroffenen Personen“. Da diese Folgen wesentlich vom Grad der Anonymisierung abhängen, muss dieser ausreichend hoch sein.

- Handelt es sich um die Erfüllung rechtlicher Pflichten, so geht der Grad der Anonymisierung zwar nicht direkt in die Bewertung der Rechtsgrundlage ein, aber die rechtlichen Pflichten können natürlich nur durch eine Anonymisierung erfüllt werden, wenn der Grad der Anonymisierung ausreichend hoch ist. Bei Verwendung dieser Rechtsgrundlage wäre im Einzelfall zu klären, ob bei einer ungenügenden Anonymisierung auch die Rechtsgrundlage nicht mehr vorliegt oder nur die relevante rechtliche Pflicht nicht erfüllt ist.

- Handelt es sich um eine Verarbeitung auf der Rechtsgrundlage des berechtigten Interesses, dann ist dafür eine Interessenabwägung erforderlich, deren Ergebnis ebenfalls wesentlich vom Grad der Anonymisierung abhängt.

In Summe ergibt sich daraus, dass eine Klärung der Rechtsgrundlage der Anonymisierung darauf eingehen muss, welcher Grad der Anonymisierung erreicht bzw. angestrebt wird. Es sollten konkrete Mindestanforderungen an eine Anonymisierung gestellt werden, die vom damit verbundenen Risiko einer Re-Identifizierung abhängen. Die wesentlichen Einflussfaktoren für die Risikobewertung, wie sie beispielsweise in einer Datenschutzfolgenabschätzung betrachtet werden müssen, sind die Höhe des ggf. eintretenden Schadens sowie die Wahrscheinlichkeit dieses Schadens. Die Schadenshöhe ist definiert durch die negativen Einflüsse auf Interessen, Grundrechte und Grundfreiheiten der betroffenen Personen, so dass man beispielsweise bei besonderen Kategorien personenbezogener Daten meist von einem höheren potentiellen Schaden und damit von höheren Anforderungen an den Grad der Anonymisierung ausgehen kann. Die Wahrscheinlichkeit des potentiellen Schadens hängt u.a. von der geplanten weiteren Verarbeitung der anonymisierten Daten ab, insbesondere der Frage, wer in welchem Umfang Zugriff auf die anonymisierten Daten erhält.

Zur Bewertung des Grades der erreichten Anonymität gibt es u.a. eine Reihe von Anonymitätsmaßen. Die bekanntesten dieser Maße sind sicher die k-Anonymität und deren Erweiterungen, wobei hier die Wahl des Parameters k eine zentrale Rolle spielt, sowie die ε-Differential Privacy mit dem Parameter ε. Ein anderer Ansatz, der kein derartiges Anonymitätsmaß verwendet, ist der Safe-Harbor-Ansatzes bei HIPAA[3], der stattdessen konkrete Anforderungen an den Umgang mit ausgewählten häufigen Identifizierern bei der Anonymisierung von Gesundheitsdaten definiert.

Ausgehend von den oben genannten Einflussfaktoren sollten vom BfDI also Empfehlungen aufgenommen werden, welcher Grad der Anonymisierung jeweils angemessen ist, mit Hinweisen zur Auswahl des Anonymitätsmaßes und den zu wählenden Parametern.

[3] Siehe `https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html`. Dieser Ansatz sollte nicht verwechselt werden mit dem gleichnamigen, mittlerweile nicht mehr relevanten, Abkommen zwischen den USA und der EU zum Datenaustausch.