Bundesministerium
des Innern, für Bau
und Heimat

Digital Gipfel
Dortmund 2019

Rolf Schwartmann / Steffen Weiß (Hrsg.)

Entwurf für einen Code of Conduct zum Einsatz DS-GVO konformer Pseudonymisierung

Ein Arbeitspapier der Fokusgruppe Datenschutz der Plattform Sicherheit, Schutz und Vertrauen für Gesellschaft und Wirtschaft im Rahmen des Digital-Gipfels 2019

Entwurf für einen Code of Conduct zum Einsatz DS-GVO konformer Pseudonymisierung der Fokusgruppe Datenschutz der Plattform Sicherheit, Schutz und Vertrauen für Gesellschaft und Wirtschaft im Rahmen des Digital-Gipfels 2019

Leitung:
Prof. Dr. Rolf Schwartmann
Kölner Forschungsstelle für Medienrecht - TH Köln

Sherpa:
Steffen Weiß, LL.M.
Gesellschaft für Datenschutz und Datensicherheit e.V.

Mitglieder:
Patrick von Braunmühl
Bundesdruckerei GmbH

Susanne Dehmel
Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V.

Philipp Ehmann
eco – Verband der Internetwirtschaft e.V.

Maximilian Hermann
Kölner Forschungsstelle für Medienrecht - TH Köln

Dr. Detlef Houdeau
Infineon Technologies AG

Angelika Hüsch-Schneider
Deutsche Telekom AG

Frank Ingenrieth, LL.M.
Selbstregulierung Informationswirtschaft e.V.

Clemens John
United Internet AG

Johannes Landvogt
Bundesbeauftragter für den Datenschutz und die Informationsfreiheit

Prof. Dr. Michael Meier
Universität Bonn/Gesellschaft für Informatik e.V.

Robin L. Mühlenbeck
Kölner Forschungsstelle für Medienrecht - TH Köln

Michael Neuber
Bundesverband Digitale Wirtschaft (BVDW) e.V.

Dr. Frank Niedermeyer
Bundesamt für Sicherheit in der Informationstechnik

Jonas Postneek
Bundesamt für Sicherheit in der Informationstechnik

Frederick Richter, LL.M.
Stiftung Datenschutz

Dr. Sachiko Scheuing
Acxiom Deutschland GmbH

Achim Schlosser
European netID Foundation

Irene Schlünder
Technologie- und Methodenplattform fürdie vernetzte medizinische Forschung e.V.

Sebastian Schulz
HÄRTING Rechtsanwälte

Dr. Tobias Stadler
Bundesbeauftragter für den Datenschutz und die Informationsfreiheit

Dr. Claus D. Ulmer
Deutsche Telekom AG

Dr. Martina Vomhof
Gesamtverband der Deutschen Versicherungswirtschaft e.V.

Benjamin Walczak
Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein

Version 1.0, 2019

Herausgeber:
Fokusgruppe Datenschutz des Digital-Gipfels

Leitung:
Prof. Dr. Rolf Schwartmann
(TH Köln/GDD)

Sherpa:
Steffen Weiß
Gesellschaft für Datenschutz und Datensicherheit e.V.,

Heinrich-Böll-Ring 10
53119 Bonn
Tel.: +49 228 96 96 75 00
info@gdd.de

Kölner Forschungsstelle für Medienrecht

Technology Arts Sciences TH Köln

Gesellschaft für Datenschutz und Datensicherheit e.V.

Vorwort

Anlässlich des Digital-Gipfels 2019 hat es sich die Fokusgruppe Datenschutz der Plattform 9 „Sicherheit, Schutz und Vertrauen für Gesellschaft und Wirtschaft“ zur Aufgabe gemacht, einen Entwurf für einen Code of Conduct für die Pseudonymisierung personenbezogener Daten zu erarbeiten. Die Pseudonymisierung weist verschiedene Bezüge zum diesjährigen Gipfelthema in Gestalt der digitalen Plattformen und der Plattformökonomie auf. Plattformen verfügen über enorme Datenmengen, mit denen u.a. KI-Anwendungen entwickelt und umgesetzt werden können. Gleichzeitig können Daten dazu verwendet werden, individuelle Profile von Nutzerinnen und Nutzern zu erstellen. Die Pseudonymisierung kann einen fundamentalen Beitrag leisten, dass Persönlichkeitsrechte von Nutzerinnen und Nutzern beim Betrieb digitaler Plattformen gewahrt werden und diese vor einer individualisierten Profilerstellung geschützt sind.

Über einen Code of Conduct für die Pseudonymisierung erhalten Betreiber von Plattformen die Möglichkeit, die Pseudonymisierung anhand transparenter Vorgaben vorzunehmen. Nutzerinnen und Nutzer profitieren von der Anwendung einheitlicher Standards. Die referenzierten Anwendungsbeispiele geben einen Einblick, in welchen weiteren Bereichen die Pseudonymisierung eine Rolle spielen kann. Das vorliegende Dokument stellt keinen finalen Code of Conduct dar. Hierzu bedarf es - neben einer Genehmigung durch eine Aufsichtsbehörde für den Datenschutz - der Festlegung von Prozessen zur Kontrolle der Einhaltung des Codes. Ebenso sollen die bestehenden Anwendungsbeispiele um sektorspezifische Good Practices erweitert werden. Denn es bedarf in der Praxis vertiefter Anschauungen, um die Ermittlung einer geeigneten Pseudonymisierungsmethode und ihre Durchführung nachvollziehen zu können. Dies wird in einer späteren Fassung des Codes erfolgen.

Allen Mitwirkenden gebührt herzlicher Dank für die kontinuierliche Arbeit in der Fokusgruppe. Besonders danke ich Herrn Rechtsanwalt Steffen Weiß von der Gesellschaft für Datenschutz und Datensicherheit für die Koordination der Arbeit.

Köln, im Oktober 2019

Professor Dr. Rolf Schwartmann

Leiter der Fokusgruppe Datenschutz der Plattform Sicherheit, Schutz und Vertrauen für Gesellschaft und Wirtschaft im Rahmen des Digital-Gipfels 2019 und Mitglied der Datenethikkommission der Bundesregierung

Inhalt

Vorwort 5

1. Einleitung 8
1.1. Anwendungsbereich 8
1.2. Begriffsdefinitionen des CoC 9

2. Prozessuale Vorgaben zum Einsatz und Betrieb einer Pseudonymisierung 9
2.1. Organisatorische Fragen 9
2.1.1. Fachverantwortlichen für den ganzen Prozess benennen 9
2.1.2. Ermittlung und Dokumentation der zur Festlegung der Pseudonymisierungsmethode notwendigen Kriterien 10
2.1.3. Risikoadäquates Rechte- und Rollenkonzept 15
2.1.4. Festlegung von Vorgaben für die Re-Identifizierung 17
2.1.5. Erfüllung von Informations- und Mitteilungspflichten gegenüber Betroffenen 17
2.1.6. Unbeabsichtigte/unrechtmäßige Aufhebung einer Pseudonymisierung 18
2.1.7. Festlegung eines Prozesses zur regelmäßigen Überprüfung der Erforderlichkeit der Verarbeitung 18
2.1.8. Mitteilungspflichten gegenüber Aufsichtsbehörden in besonderen Fällen 19
2.1.9. Dokumentation und regelmäßige Evaluation des Prozesses, der erfolgten Abwägungen und der tatsächlich getroffenen Maßnahmen 19
2.2. Technische Fragen 20
2.2.1. Allgemeine Anforderungen an die Pseudonymisierung 20
2.2.2. Allgemeine Anforderungen an Identifikatoren (IDs) 21
2.2.3. Berechnungsverfahren 21

3. Anwendungsbeispiele der Pseudonymisierung 23

1. Einleitung

Ziel dieses Code of Conduct (CoC) ist es, entsprechend Art. 40 Abs. 2 lit. d Datenschutz-Grundverordnung (DS-GVO) konkrete Verhaltensregeln für eine datenschutzkonforme Pseudonymisierung nach den Anforderungen der DS-GVO zu beschreiben.

Die Pseudonymisierung schützt betroffene Personen vor einer ungewollten Identifikation und ist eine Umsetzung des Grundsatzes der Datensparsamkeit aus Art. 5 Abs. 1 lit. b DS-GVO. Sie stellt eine technisch-organisatorische Schutzmaßnahme nach Maßgabe der Art. 25, 32 DS-GVO dar. Gleichwohl beeinflusst sie auch die Rechtmäßigkeit einer Verarbeitung personenbezogener Daten, wie etwa Art. 6 Abs. 4 lit. e DS-GVO zeigt.
Sie erfüllt damit sowohl eine Schutz- als auch eine Ermöglichungsfunktion. Die Pseudonymisierung zeichnet sich nach ihrer gesetzlichen Definition dadurch aus, dass personenbezogene Daten in einer Weise verarbeitet werden, dass diese Daten ohne Hinzuziehung zusätzlicher Informationen nicht mehr einer spezifischen Person zugeordnet werden können (vgl. Art. 4 Nr. 7 DS-GVO).

D.h. ein unmittelbarer Personenbezug ist im Rahmen einer Pseudonymisierung möglich, muss jedoch abseits einer gewollten Aufdeckung mittels technisch-organisatorischer Maßnahmen verhindert werden. Die DS-GVO enthält weder technisch-organisatorische Hinweise darüber, wie ein Pseudonym erstellt werden kann, noch macht sie Angaben zu möglichen Schutzmaßnahmen bezüglich des erstellten Pseudonyms.
Zu diesem Zweck werden in diesem Code of Conduct sowohl prozessuale als auch organisatorische und technische Vorgaben definiert, die sowohl Verantwortlichen als auch Auftragsverarbeitern eine praxisnahe Umsetzung der Pseudonymisierung ermöglichen.

1.1. Anwendungsbereich

Dieser CoC gilt für Verantwortliche oder Auftragsverarbeiter unabhängig ihrer Branche oder ihres Sektors, wenn sie personenbezogene Daten nach den Anforderungen der DS-GVO selbst pseudonymisieren oder die Anwendung der Pseudonymisierung personenbezogener Daten verantworten. Die Ausführungen des CoC gelten unabhängig von der internen Organisations- und Aufgabenverteilung des Verantwortlichen oder Auftragsverarbeiters.

Verantwortliche oder Auftragsverarbeiter, die in ihren Diensten oder Produkten pseudonymisierte Daten einsetzen, können diesem CoC beitreten, um nachzuweisen, dass die verwendeten Pseudonyme nach den hierin definierten Regeln erstellt wurden.

Verantwortliche und Auftragsverarbeiter werden in der Regel sowohl Datenverarbeitungen betreiben, die in Zusammenhang mit einer Pseudonymisierung stehen, als auch solche, die in keinerlei Bezug zu einer Pseudonymisierung stehen. Selbst soweit Datenverarbeitungen in Verbindung mit Pseudonymisierung stattfinden, ist davon auszugehen, dass insbesondere bei international tätigen Verantwortlichen oder Auftragsverarbeitern, nicht jede Datenverarbeitung der DS-GVO unterfällt oder diesem CoC unterworfen werden soll. Insofern können Verantwortliche und Auftragsverarbeiter selbst entscheiden, welche Pseudonymisierungsprozesse diesem CoC unterworfen werden. Bei denjenigen Produkten, Dienstleistungen oder sonstigen Datenverarbeitungen, die auf Pseudonyme zurückgreifen, die Pseudonymisierungsprozessen entstammen, die diesem CoC unterworfen waren, ist auf diesen Umstand transparent hinzuweisen.

1.2. Begriffsdefinitionen des CoC

- **Pseudonymisierung** bedeutet Pseudonymisierung i.S.d. Art. 4 Nr. 5 DS-GVO: Pseudonymisierung [ist] die Verarbeitung personenbezogener Daten in einer Weise, dass die personenbezogenen Daten ohne Hinzuziehung zusätzlicher Informationen nicht mehr einer spezifischen betroffenen Person zugeordnet werden können, sofern diese zusätzlichen Informationen gesondert aufbewahrt werden und technischen und organisatorischen Maßnahmen unterliegen, die gewährleisten, dass die personenbezogenen Daten nicht einer identifizierten oder identifizierbaren natürlichen Person zugewiesen werden.

- Nach Art. 4 Nr. 5 DS-GVO ist die **zusätzliche Information** die einzige Information, mit der die Verbindung eines Pseudonyms zu der repräsentierten Person hergestellt werden kann. Abhängig von der Pseudonymisierungsmethode kann die zusätzliche Information eine direkte Zuordnung oder eine Zuordnungsregel sein.

- Ein **Pseudonym** ist eine Zeichenkette, die Identitätsdaten einer Person ersetzt und damit diese Person repräsentiert.

- **Pseudonymisierungsmethode** bezeichnet das technisch-organisatorische Verfahren, mit dem ein Pseudonym generiert wird.

- **Fachverantwortliche** sind alle Personen oder Abteilungen innerhalb eines Unternehmens oder einer öffentlichen Stelle, die nicht für die Organisation der gesamten Verarbeitungstätigkeit zuständig sind, sondern nur einzelne Teilbereiche datenschutzkonform ausgestalten (wie etwa die ordnungsgemäße Pseudonymisierung von personenbezogenen Daten).

- **Fachverantwortliche für Pseudonymisierung** (FvfP) sind alle Personen oder Abteilungen innerhalb eines Unternehmens oder einer öffentlichen Stelle, die sich für die datenschutzkonforme Ausgestaltung des Pseudonymisierungsprozesses übergeordnet, jedenfalls in Form einer übergeordneten Aufsichts- und Beratungsinstanz, verantwortlich zeichnen.

2. Prozessuale Vorgaben zum Einsatz und Betrieb einer Pseudonymisierung

2.1. Organisatorische Fragen

2.1.1. Fachverantwortlichen für den ganzen Prozess benennen

In organisatorischer Hinsicht ist seitens des oder der Verantwortlichen bzw. des Auftragsverarbeiters ein Fachverantwortlicher für Pseudonymisierung (FvfP) zu benennen. Die in der DS-GVO festgelegten Aufgaben und Pflichten des Verantwortlichen werden nicht übertragen. Dieser FvfP koordiniert die einzelnen organisatorischen Verantwortlichkeiten vor, während und nach der Durchführung der Pseudonymisierung.

Erläuterung: Hierbei meint der Begriff des FvfP nicht den datenschutzrechtlich Verantwortlichen im Sinne der DS-GVO, sondern (untechnisch) den für die Organisation und den ordnungsgemäßen Ablauf der Pseudonymisierung intern Verantwortlichen. Die Pseudonymisierung personenbezogener Daten ist in der Regel Teil einer allgemeineren Verarbeitungstätigkeit (gemäß Verarbeitungsverzeichnis).

Der FvfP kann durchaus auch andere Fachverantwortlichkeiten übernehmen oder auch die Gesamtverantwortung für die jeweilige Datenverarbeitung tragen. Es ist aber in jedem Fall ungeachtet anderer Zuständigkeiten die Zuständigkeit dieser Person oder Abteilung als FvfP zu dokumentieren. Die Benennung eines Datenschutzbeauftragten zum FvfP ist nicht zulässig.

Erläuterung: Der FvfP ist nicht gleichzusetzen mit dem Datenschutzbeauftragten des Verantwortlichen oder Auftragsverarbeiters. Im Gegensatz zum FvfP trägt der Datenschutzbeauftragte keine Verantwortung für die Gesetzmäßigkeit einer Datenverarbeitung. Seine/ihre gesetzlichen Aufgaben sind in Art. 39 DS-GVO definiert und zeichnen sich durch die Beratung und Kontrolle aus. Im Bereich der Pseudonymisierung kann der Datenschutzbeauftragte sowohl bei ihrer Planung und Durchführung beraten als auch die Einhaltung der gesetzlichen Anforderungen an die Pseudonymisierung sowie dieses CoC kontrollieren. Aufgrund der Zuweisung einer organisatorischen Verantwortlichkeit für den FvfP wäre eine Identität von Datenschutzbeauftragtem und FvfP nicht mit den gesetzlichen Anforderungen in Einklang zu bringen.

Der FvfP muss das für die Pseudonymisierung erforderliche technische und organisatorische Fachwissen besitzen. Soweit als FvfP eine Abteilung benannt wurde, muss die Abteilung in (Teil-)Gesamtheit das nötige Fachwissen aufweisen, wenn und soweit organisatorisch sichergestellt ist, dass diese Abteilung die Verantwortlichkeit stets in entsprechender (Teil-)Gesamtheit ausübt.

2.1.2. Ermittlung und Dokumentation der zur Festlegung der Pseudonymisierungsmethode notwendigen Kriterien

Für den rechtskonformen Einsatz der Pseudonymisierung sind die nachstehenden Kriterien in dokumentierter Form zu berücksichtigen.

2.1.2.1. Art und Risikoklasse der verarbeiteten personenbezogenen Daten

Zur Gewährleistung einer datenschutzkonformen Pseudonymisierung ist die Art und Risikoklasse der verarbeiteten Daten festzulegen. Auf Basis dieser Risikoabschätzung hat die Auswahl des adäquaten, DS-GVO konformen Pseudonymisierungsverfahrens zu erfolgen.

Erläuterung: Grundsätzlich kann es verschiedene Kategorien von Daten geben:
- Personenbezogene Daten nach Art. 4 Nr. 1 DS-GVO
- Besondere Kategorien personenbezogener Daten nach Art. 9 Abs. 1 DS-GVO

Die verarbeiteten Datenkategorien können dem Verzeichnis der Verarbeitungstätigkeiten entnommen werden. Im Rahmen der Risikobewertung der verarbeiteten Daten können zudem etwa Einschätzungen aus Risikoanalysen oder einer Datenschutz-Folgenabschätzung zur Anwendung kommen.

Die verwendete Datenkategorie stellt jedoch für sich kein taugliches Kriterium für eine Risikoeinschätzung dar und kann allenfalls als Indiz herangezogen werden. Vielmehr müssen auch andere Aspekte im Rahmen der Risikoeinschätzung Berücksichtigung finden. Dies ist z.B.
- der Zweck und der Kontext der Verarbeitung (s. nachfolgend 2.1.2.2. und 2.1.2.3.); So können beispielsweise die identischen personenbezogenen Daten im Rahmen einer Vertragserfüllung oder zur Verfolgung von Nutzeraktivitäten verwendet werden;
- die Kategorie der Betroffenen; z.B. Kinder oder Angehörige bestimmter Bevölkerungsgruppen ohne sogleich den Anwendungsbereich des Art. 9 Abs. 1 DS-GVO auszulösen;
- Zahl der Betroffenen (s. nachfolgend 2.1.2.4.) oder die Kombination der unterschiedlichen Datenkategorien.

2.1.2.2. Beabsichtigte Verarbeitungszwecke

Es ist festzuhalten, zu welchen Zwecken die Daten verarbeitet werden sollen.

Erläuterung: Es kann mehr als einen Zweck der Verarbeitung geben. Zwecke können nicht ohne Weiteres im Nachgang einer Datenerhebung geändert werden, so dass diese durchaus möglichst umfassend dokumentiert werden sollten. Zwecke müssen aber auch hinreichend präzise sein, sodass die Beachtung des Zweckbindungsgrundsatzes ermöglicht wird. In Betracht kommen etwa Datenverarbeitungen zu Abrechnungszwecken, zur Prüfung der Netzauslastung eines Mobilfunkanbieters, für Zwecke der Produktentwicklung oder die Verarbeitung von Daten zu Forschungszwecken. Forschungszwecke sollten hierbei insoweit im Rahmen der Dokumentation präzisiert werden, dass der Forschungskontext oder das Forschungsziel tatsächlich dergestalt nachvollzogen werden kann, ob eine tatsächliche, zukünftige Verarbeitung dem intendierten Forschungszweck unterfällt und somit auch eine Risikoeinschätzung hinreichend abgeleitet werden kann. Die Beschreibung des Zwecks hat auch Einfluss auf die datenschutzrechtliche Beurteilung, ob die Datenverarbeitung zu den angestrebten Zwecken noch unter den einschlägigen Erlaubnistatbestand fällt und zum anderen ist zu prüfen, ob die Pseudonymisierung etwas an dieser Beurteilung ändert.

2.1.2.3. Kontext der Pseudonymisierung

Der Kontext der Pseudonymisierung ist zu dokumentieren.

Erläuterung: Mit dem Kontext der Verarbeitung ist der rechtliche Kontext für die Pseudonymisierung gemeint. Eine Pseudonymisierung kann beispielsweise im Zuge ihrer Ermöglichungsfunktion im Rahmen von Art. 6 Abs. 1 lit. f sowie Art. 6 Abs. 4 DS-GVO oder als rein technisch-organisatorische Maßnahmen nach Art. 32 DS-GVO bzw. im Rahmen des Art. 25 DS-GVO eingesetzt werden.

Eine Dokumentation ist erforderlich, da dieser Kontext ebenfalls Einfluss auf die Wahl des angemessenen Pseudonymisierungsverfahrens hat.

2.1.2.4. Erwartete Anzahl der verarbeiteten Datensätze

Es ist zu prüfen und zu dokumentieren, wie hoch die Anzahl der verarbeiteten Datensätze sein wird.

Erläuterung: Es muss ein Überblick darüber bestehen, ob nur wenige Datensätze oder eine große Anzahl von Daten pseudonymisiert werden. Im Rahmen der Prüfung der Zahl der zu verarbeitenden Datensätze ist relevant, ob die Datensätze statisch oder dynamisch sind, also ob es sich um eine festgelegte Anzahl von Daten handelt, die pseudonymisiert wird oder ob der Datensatz fortlaufend durch weitere Daten angereichert wird. Klassische Listenverfahren zur Pseudonymisierung eignen sich beispielsweise nicht für eine große Anzahl von Daten.

2.1.2.5. Geeignete Pseudonymisierungsarten

Die Arten der benötigten Pseudonyme sind zu dokumentieren.

Erläuterung: Unterschiedliche Arten von Pseudonymen eigen sich beispielsweise für bestimmte Einsatzzwecke besonders obschon diese für andere Einsatzzwecke gänzlich ungeeignet sein können. Es kann zwischen den folgenden Pseudonymisierungsarten unterschieden werden:

- Personen-Pseudonyme, die an Stelle von Identitätsdaten wie z.B. Name, Ausweisnummer oder Mobiltelefonnummer stehen
- Rollen-Pseudonyme, bei denen eine oder ggf. mehrere Personen einem Pseudonym zugeordnet sind (z.B. IP-Nummer)
- Beziehungs-Pseudonyme, bei denen eine Person für jede (Kommunikations-) Beziehung ein anderes Pseudonym verwendet, z.B. unterschiedliche Spitznamen
- Rollen-Beziehungs-Pseudonyme, die eine Kombination der beiden Pseudonym-Arten sind
- Wechselnde Pseudonyme, bei denen z.B. für jede Transaktion oder jeden Eintrag ein neues Pseudonym genutzt wird, was z.B. beim Online-Banking zum Einsatz kommt

Es sind unter Berücksichtigung des Zwecks und des Kontextes der Verarbeitung solche Arten von Pseudonymen vorzuziehen, die für den jeweiligen Einsatzzweck geeignet sind und den Betroffenen gleichzeitig im größtmöglichen Umfang vor einer ungewollten Identifizierung schützen. Der FvfP unterstützt bei der Auswahl der geeigneten Pseudonymisierungsart. Die zur Entscheidung für bzw. gegen eine in Betracht kommende Pseudonymisierungsart erfolgte Abwägung ist zu dokumentieren.

Erläuterung: Generell ist das Risiko einer Aufdeckung von Personen-Pseudonymen höher als von Rollen- bzw. Beziehungs-Pseudonymen. Dies steht im Zusammenhang mit der Verknüpfung eines Pseudonyms mit einer dahinterstehenden Person. Je nach Zweck und Kontext der Verarbeitung kann die Verwendung von Personen-Pseudonymen erforderlich sein. Andererseits besteht ein geringeres Risiko der Aufdeckung von Personen bei Rollen-Beziehungs-Pseudonymen und wechselnden Pseudonymen als bei den erwähnten Personen-Pseudonymen.

2.1.2.6. Festlegung der geeigneten Pseudonymisierungsmethode und des Zeitpunkts der Pseudonymisierung

Für die Pseudonymisierung stehen unterschiedliche Methoden zur Verfügung[1].

Die Stärke der angewandten Methode muss unter Berücksichtigung aller objektiven Faktoren, Risiken für die Rechte und Freiheiten der Betroffenen sowie auch den Kosten der Identifizierung und der dafür erforderliche Zeitaufwand bei Einsatz der zum Zeitpunkt der Verarbeitung verfügbaren Technologien sowie absehbaren technologischen Entwicklungen, geprüft und entsprechend festgelegt und dokumentiert werden. Beim Einsatz von Berechnungsverfahren ist ein State-of-the-Art-Transformationsverfahren einzusetzen (zu den technischen Anforderungen vgl. 2.2.1.).

Pseudonymisierungsverfahren sind im Übrigen so auszugestalten, dass eine einfache und effiziente Selektion und Löschung der Daten möglich ist, soweit der Verarbeitungszweck nicht mehr besteht oder für die Verarbeitung keine Rechtsgrundlage mehr besteht.

Erläuterung: Der Grundsatz der Datenminimierung ist stets zu befolgen. Ebenfalls sind die Grundsätze des Privacy-by-Design zu berücksichtigen. Demzufolge hat die technische Ausgestaltung von Anfang an die entsprechenden Rahmenbedingungen bereitzustellen. Die Einhaltung dieser Grundsätze vermeidet somit die an sich unzulässige fortwährende Speicherung von schwer re-identifizierbaren, pseudonymisierten Daten soweit. Ergänzend sind solche Pseudonymisierungsverfahren zu bevorzugen, die eine nachträgliche Anonymisierung von Daten einfach ermöglichen.

Die Pseudonymisierung hat im Verarbeitungsprozess so früh wie möglich zu erfolgen.

Erläuterung: Personenbezogene Daten müssen auf das für die Zwecke der Verarbeitung notwendige Maß beschränkt sein (vgl. Art. 5 Abs. 1 lit. c DS-GVO). Wenn die Pseudonymisierung als geeignete Verarbeitung durch den Verantwortlichen oder Auftragsverarbeiter identifiziert worden ist, sollte ihre technische Umsetzung zeitnah durchgeführt werden. Ebenso sollte auch in mehrstufigen Datenverarbeitungen die Pseudonymisierung möglichst frühzeitig vorgenommen werden, insbesondere wenn nicht-pseudonymisierte Daten auf den vorgelagerten Verarbeitungsstufen nicht erforderlich sind.

2.1.2.7. Geplante Weitergabe der pseudonymisierten Daten

Es ist zu dokumentieren, ob pseudonymisierte Daten an Dritte übermittelt werden sollen. Der Verantwortliche oder Auftragsverarbeiter hat angemessene Maßnahmen zu treffen, dass die weitergegebenen Daten durch den oder die Empfänger nur zu den zuvor bestimmten Zwecken verarbeitet werden. Der Verantwortliche oder Auftragsverarbeiter hat sicherzustellen, dass

[1] Schwartmann/Weiß (Hrsg.), Anforderungen an den datenschutzkonformen Einsatz von Pseudonymisierungslösungen - Ein Arbeitspapier der Fokusgruppe Datenschutz der Plattform Sicherheit, Schutz und Vertrauen für Gesellschaft und Wirtschaft im Rahmen des Digital-Gipfels 2018, D.2.2 ff.

die Weitergabe der pseudonymisierten Daten an den Empfänger von einer Rechtsgrundlage abgedeckt ist. Darüber hinaus hat der Verantwortliche oder der Auftragsverarbeiter angemessene Maßnahmen zu treffen, um eine unzulässige Re-Identifizierung von Betroffenen durch den Empfänger möglichst auszuschließen.

Erläuterung: Da auch pseudonymisierte Daten einen Personenbezug haben, gelten die allgemeinen datenschutzrechtlichen Anforderungen an die Verarbeitung, so auch die Zweckbindung gem. Art. 5 Abs. 1 lit. b DS-GVO. Die datenabgebende Stelle sowie der Empfänger haben sich vor der Weitergabe pseudonymisierter Daten entsprechend auf einen Zweck zu einigen. Als angemessene Maßnahme zur Sicherstellung kann beispielsweise der Verarbeitungszweck seitens des Empfängers in Textform oder schriftlich (z.B. als Bestandteil eines Vertrages) bestätigt werden. Da lediglich die pseudonymisierte Datenweitergabe in den Geltungsbereich dieses CoC fällt, hat eine Identifizierung von Betroffenen im Rahmen der Datenweitergabe zu unterbleiben. Der CoC formuliert daher eine diesbezügliche Pflicht angemessener Sicherungsmaßnahmen der datenabgebenden Stelle vor der Datenweitergabe an den Empfänger. Dies umfasst beispielsweise eine Prüfpflicht hinsichtlich offensichtlicher Identifizierungsmöglichkeiten durch den Empfänger, da eine detaillierte Kenntnis von Verknüpfungsmöglichkeiten der Daten auf Empfängerseite nicht vorausgesetzt werden kann. Der Verantwortliche oder Auftragsverarbeiter sollte sich in jedem Fall ergänzend in Textform oder schriftlich (z.B. als Bestandteil eines Vertrages) bestätigen lassen, dass eine Identifikation durch den Empfänger nicht stattfindet. Soweit die Prüfung ergeben hat, dass der Empfänger offensichtlich eine Re-Identifizierung durchführen könnte, sollten, soweit dies aufgrund der zu erwartenden Risiken für Betroffene erforderlich erscheint, angemessene, ergänzende Schutzmaßnahmen implementiert werden.

Hinsichtlich der Rechtsgrundlage für die Weitergabe muss insbesondere in Fällen, in denen eine Einwilligung für die Erhebung personenbezogener Daten eingeholt wurde, der Umstand der Weitergabe an andere Stelle in pseudonymisierter Form hiervon erfasst sein. Andernfalls bedarf es einer weiteren Rechtsgrundlage.

2.1.2.8. Geplante Verarbeitung der pseudonymisierten Daten im Drittstaat

Es ist zu dokumentieren, ob pseudonymisierte Daten außerhalb des EWR verarbeitet werden sollen. Für den Fall der Übermittlung personenbezogener Daten an ein Drittland hat der Verantwortliche oder Auftragsverarbeiter dafür zu sorgen, dass die Vorgaben des Kapitels V der DS-GVO bezüglich der Gewährleistung eines angemessenen Datenschutzniveaus eingehalten werden. Der Betroffene ist im Rahmen der Erhebung seiner personenbezogenen Daten auf die Übermittlung dieser Daten in ein Drittland hinzuweisen, auch wenn ausschließlich pseudonymisierte Daten übermittelt werden.

Erläuterung: Die DS-GVO stellt besondere Anforderungen an die Verarbeitung personenbezogener Daten in einem Drittland außerhalb der EU bzw. des EWR. Diese Anforderungen sind in Kapitel V der DS-GVO geregelt und beinhalten beispielsweise die Übermittlung personenbezogener Daten auf Basis eines Angemessenheitsbeschlusses der EU-Kommission oder anderer geeigneter Garantien nach Art. 46 -GVO. Der Umstand, dass die zu übermittelnden Daten pseudonymisiert sind, soll die datenabgebende Stelle nicht davon befreien, die Anforderungen des Kapitels V einzuhalten. Immerhin kann ein Betroffener auch im Drittland unter Verwendung des Schlüssels zur Pseudonymisierung re-identifiziert werden.

2.1.2.9. Geplante/absehbare Häufigkeit der Re-Identifizierung?

Die geplante oder absehbare Häufigkeit der Re-Identifizierung von Betroffenen ist in dokumentierter Form festzulegen.

Erläuterung: Die gewählten Verarbeitungszwecke haben Einfluss auf die Frage, ob zeitnah und kurzfristig eine Re-Identifizierung von Betroffenen vorgenommen werden muss. Beispielsweise kann es im Bereich der Netzwerküberwachung auf Basis von Pseudonymen kurzfristig notwendig sein, einen mit einem Schadcode infizierten Arbeitsplatz zu ermitteln.

Die geplante bzw. erwartete Häufigkeit der Re-Identifizierung von Datensätzen ist zu definieren. Hierbei ist auch zu dokumentieren, aus welchen geplanten bzw. erwarteten Gründen bzw. zu welchen Zwecken eine solche Re-Identifikation erfolgen wird (z.B. Zur Wahrung der Rechte der Betroffenen). Neben den Gründen und Zwecken ist zudem zu dokumentieren, welche Verzögerungstoleranz im Falle einer Re-Identifizierung besteht, d.h. wie hoch die maximale Verzögerung bis zur hinreichenden Re-Identifikation eines Datensatzes sein darf.

Erläuterung: Pseudonymisierungsmethoden unterscheiden sich unter anderem in ihrer Effizienz und Handhabbarkeit bezüglich durchzuführender Re-Identifikationen. Ebenso interagiert die Häufigkeit der erwarteten Re-Identifikationen auch mit einer angemessenen Aufteilung von Funktionen im Sinne des Abschnitts 2.1.3. Die hier zu erstellende Dokumentation soll es dem FvfP ermöglichen, einerseits eine für sich verbindliche Abwägungsgrundlage zu schaffen. Andererseits soll es dem FvfP ermöglichen, über die Zeit die hier aufgestellte Hypothese und Planung zu evaluieren. Bei einer solchen Evaluation wäre z.B. zu berücksichtigen, ob eine eventuell sehr hohe, erwartete Re-Identifizierungsquote tatsächlich in der Praxis erfolgt. Ebenso wäre z.B. zu berücksichtigen, ob die Verzögerungstoleranz eingehalten werden konnte oder ob durch neue technische Entwicklungen inzwischen auch andere Pseudonymisierungsmethoden in der Lage sind, diese Toleranzen einzuhalten.

2.1.3. Risikoadäquates Rechte- und Rollenkonzept

Hinsichtlich des Zugriffs auf pseudonymisierte Daten und deren für die jeweilige Tätigkeit erforderlichen Kombinationen, möglicherweise bestehende Übersetzungstabellen und Schlüssel zur Re-Identifizierung einer Person und sonstige eine Aufhebung des Pseudonyms fördernde Informationen ist ein angemessenes Rechte- und Rollenkonzept vorzusehen. Je sensibler die verarbeiteten Daten bzw. je höher die zu erwartenden Risiken für die Rechte und Freiheiten der Betroffenen, umso effektiver ist eine solche Trennung auszugestalten.

Erläuterung: Ausweislich der gesetzlichen Definition der Pseudonymisierung müssen zusätzliche Informationen, die eine Identifikation von Betroffenen ermöglichen, gesondert aufbewahrt und eine Identifikation durch technisch-organisatorische Maßnahmen verhindert werden. Ein bestehendes Rechte- und Rollenkonzept kann eine solche technisch-organisatorische Maßnahme darstellen. Innerhalb eines solchen Rechte- und Rollenkonzepts eignen sich - je nach Risiko der Daten und dem Kontext der Verarbeitung - unterschiedliche Modelle:

"Alles-in-einer-Hand"-Modell: Hier verfügt der Verantwortliche oder Auftragsverarbeiter sowohl über die pseudonymisierten Daten als auch über die jederzeitige Möglichkeit, die verarbeiteten Pseudonyme aufzuheben bzw. Betroffene zu re-identifizieren. Hierbei kann eine Re-Identifizierungsmöglichkeit bei einer Person, einer Abteilung oder in einer juristischen Person verankert sein. In diesen Fällen sollten zumindest interne Vorgaben existieren, aus denen sich zulässige und unzulässige Umstände zur Durchführung einer Re-Identifizierung sowie etwaige Dokumentationspflichten über erfolgte Re-Identifikationen ergeben. Mit Anstieg der zu erwartenden Risiken sollten diese internen Vorgaben zudem um ein angemessenes, internes Rechte- und Rollenkonzept nach dem Need-to-know-Prinzip ergänzt werden (vgl. Mischmodelle).

Treuhändermodell: Im klassischen Treuhändermodell ist der Treuhänder eine juristische Person außerhalb des Verantwortlichen oder Auftragsverarbeiters, mithin ein "Dritter". Er ist somit eine von der Datenerhebung und Datenauswertung räumlich und organisatorisch unabhängige Vertrauensstelle. Ein Treuhänder kann beispielsweise mit der Aufbewahrung von Schlüsseln zur Re-Identifizierung von Betroffenen betraut werden. Ebenso ist die Verarbeitung von Pseudonymen durch ihn denkbar, während etwaige Schlüssel und Rohdaten beim Verantwortlichen bzw. Auftragsverarbeiter verbleiben.

Die Schlüsselverwaltung ist der voraussichtliche Regelfall, bei welchem ein Treuhänder auf verschiedene Art und Weise eingebunden werden kann. Auch die innerhalb des Treuhändermodells gewählte Art und Weise sollte sich stets an den dokumentierten Risiken für die Betroffenen orientieren:

- Ex ante: Der Treuhänder führt eine Re-Identifizierung von Betroffenen zu bereits vor Beginn der Verarbeitung definierten Zwecken und zuvor definierten Sachverhalten durch.
- Ad hoc: Der Treuhänder führt eine Re-Identifizierung von Betroffenen auf Basis zuvor definierter Abwägungskriterien, aber im Rahmen nicht zuvor definierter Zwecke und Sachverhalte durch.
- Ex post: Der Treuhänder wird über erfolgte Re-Identifizierungen samt Grund informiert (z.B. als Einzelfall oder über Statistiken). Der Treuhänder kann diese Informationen evaluieren und auf deren Basis angemessene Maßnahmen treffen; z.B. Schulungen oder auch Disziplinarmaßnahmen.

Mischmodelle: Denkbar sind auch Mischmodelle. Hierbei kann z.B. die Trennung der zur Re-Identifikation notwendigen Informationen auch innerhalb des Verantwortlichen bzw. Auftragsverarbeiters erfolgen, in dem die Informationen einem Rechte- und Rollenkonzept unterworfen werden. Dies kann z.B. auch umfassen, dass Informationen über mehrere Hierarchie-Ebenen oder an sich unabhängigen Abteilungen verteilt werden. Hierzu könnten sich auch ohnehin der Organisation für derartige Fragestellungen zuständige Abteilungen (Innenrevision bzw. Compliance oder Rechtabteilung, (IT-)Sicherheits- oder Datenschutzbeauftragte) eignen. Gerade in großen Organisationen bietet sich aber auch der Aufbau einer eigenen vertrauenswürdigen „dritten Partei" an, welche die getrennte Verwaltung der Daten und/oder Geheimnisse bzw. Schlüssel intern anbietet.

Derartige Mischmodelle sind beispielsweise insbesondere in den Fällen denkbar, in denen die Verarbeitung mehrere Verarbeitungsschritte und mehrere Pseudonymisierungsstufen umfasst, für welche jeweils unterschiedliche Risiken für die Betroffenen dokumentiert wurden.

2.1.4. Festlegung von Vorgaben für die Re-Identifizierung

Für den Fall, dass eine Re-Identifizierung von Betroffenen auf Basis der pseudonymisierten Daten vorgesehen ist, sind folgende Vorgaben zu beachten und deren Prüfung zu dokumentieren. Der FvfP unterstützt hierbei:

1. Bei der Pseudonymisierung als reine Schutzmaßnahme bedarf es über die ursprüngliche Legitimation zur Datenverarbeitung hinaus keiner Erlaubnis zur Rückführung der Pseudonyme auf Einzelpersonen. Die Rückführung ist vom ursprünglichen Verwendungszweck gedeckt.
2. Bei einer Pseudonymisierung zur Ermöglichung der Weiterverarbeitung von Daten nach Art. 6 Abs. 4 DS-GVO gilt folgendes:
   - In Fällen, in denen der Betroffene ein überwiegendes Interesse an der Rückführung hat (z.B. zum Zweck einer Information oder einer Widerspruchsmöglichkeit), ist die Zulässigkeit in Abhängigkeit der verarbeiteten Daten zu prüfen (Art. 6 bzw. Art. 9 DS-GVO).
   - In Fällen, in denen nicht festgestellt werden kann, ob der Betroffene ein Interesse an der Rückführung hat, ist eine Einwilligung zur Re-Identifizierung einzuholen. Hiervon ausgenommen sind Rückführungen auf Basis einer rechtlichen Erlaubnis.
   - In Fällen, in denen der Verantwortliche ein überwiegendes Interesse an der Rückführung (z.B. zum Zweck einer Information) hat, bedarf es der Prüfung der Zulässigkeit in Abhängigkeit der verarbeiteten Daten (Art. 6 Abs. 4 DS-GVO)
3. In Fällen, in denen ein dynamischer Datensatz (vgl. 2.1.2.4) pseudonymisiert wird, ist in regelmäßigen Abständen zu prüfen, ob durch diese Dynamik eine Re-Identifizierung von Betroffenen möglich wird. Im Falle der Möglichkeit einer Re-Identifizierung gelten die Vorgaben der Ziffern 1 und 2.

2.1.5. Erfüllung von Informations- und Mitteilungspflichten gegenüber Betroffenen

Wenn die Pseudonymisierung nur als technisch-organisatorische Maßnahme mit Schutzfunktion eingesetzt wird, ist keine gesonderte Information über die allgemeinen Datenschutzhinweise hinaus erforderlich.
Soll eine Weiterverarbeitung zu kompatiblen Zwecken erfolgen, sind folgende Zwe-

cke zu unterscheiden:
1. Die kompatible Weiterverarbeitung gemäß Art. 6 Abs. 4 DS-GVO ist von Anfang an beabsichtigt – dann sollte die Information direkt in den Datenschutzhinweisen erfolgen.
2. Die kompatible Weiterverarbeitung wird erst zu einem späteren Zeitpunkt entschieden - dann ist zu diesem Zeitpunkt eine Information der Betroffenen gemäß Art. 13 Abs. 3 erforderlich.

Die Informations- und Mitteilungspflichten gegenüber den Betroffenen beziehen sich auch auf etwaige Widerrufsbefugnisse oder Einwilligungserfordernisse.

2.1.6. Unbeabsichtigte/unrechtmäßige Aufhebung einer Pseudonymisierung

Für den Fall einer unbeabsichtigten oder unrechtmäßigen Aufhebung einer Pseudonymisierung ist ein Reaktionsplan festzulegen. Der FvfP unterstützt hierbei. Der Reaktionsplan hat folgende Aspekte zu umfassen:

- Bewertung des Risikos für Betroffene
- Maßnahmen zur Abwendung/Eindämmung des Risikos
- Bewertung einer Meldepflicht nach Art. 33/Art. 34 DS-GVO
- Meldung an die Aufsichtsbehörde und den Betroffenen bei Bestehen einer Meldepflicht.

Der Reaktionsplan kann in einen bestehenden Prozess (z.B. Incident-Response-Plan) beim Verantwortlichen oder Auftragsverarbeiter eingebunden werden.

Erläuterung: Die Aufhebung einer Pseudonymisierung kann gem. ErwG 85 S. 1 eine Datenschutzverletzung darstellen, die im Falle eines mit der Verletzung verbundenen Risikos für den Betroffenen einer Aufsichtsbehörde bzw. im Falle eines voraussichtlich hohen Risikos auch an den Betroffenen zu melden ist. Verantwortliche und Auftragsverarbeiter sollen daher notendige Schritte im Falle der Aufhebung einer Pseudonymisierung in einem Reaktionsplan festhalten. Der Reaktionsplan muss dabei nicht gesondert für die Pseudonymisierung erstellt werden, sondern kann allgemein für Datenschutzvorfälle beim Verantwortlichen oder Auftragsverarbeiter bestehen, muss jedoch die Aufhebung einer Pseudonymisierung ausdrücklich adressieren.

2.1.7. Festlegung eines Prozesses zur regelmäßigen Überprüfung der Erforderlichkeit der Verarbeitung

Es ist durch zu definieren und zu dokumentieren, in welchen Abständen die Erforderlichkeit der Verarbeitung der pseudonymisierten Daten zu überprüfen ist. Der FvfP berät und unterstützt hierbei. Eine solche Überprüfung sollte in der Regel mindestens alle zwei Jahre erfolgen. Die Überprüfung ist zu dokumentieren. Wird im Rahmen dieser Überprüfung festgestellt, dass die Verarbeitung nicht mehr erforderlich ist, sind die pseudonymisierten Daten datenschutzkonform zu löschen oder zu anonymisieren.

Erläuterung: Da pseudonymisierte Daten eine Re-Identifizierung von Betroffenen ermöglichen, unterliegt auch solch eine Verarbeitungstätigkeit dem Prinzip der Speicherbegrenzung aus Art. 5 Abs. 1 lit. e DS-GVO. Werden pseudonymisierte Daten für den festgelegten Zweck der Verarbeitung nicht mehr benötigt, sind sie zu löschen. Folglich ist ein Turnus für eine Prüfung der Erforderlichkeit durch den Verantwortlichen oder Auftragsverarbeiter festzulegen, um die Erforderlichkeit der Verarbeitung festzustellen.

2.1.8. Mitteilungspflichten gegenüber Aufsichtsbehörden in besonderen Fällen

Ist trotz einer Pseudonymisierung weiterhin ein hohes Risiko für Rechte und Freiheiten Betroffener[2] im Rahmen einer Verarbeitungstätigkeit feststellbar und stellt die Pseudonymisierung die einzige Schutzmaßnahme dar, ist die zuständige Aufsichtsbehörde gem. Art. 36 DS-GVO zu konsultieren. Der FvfP muss hierbei beratend hinzugezogen werden.

Erläuterung: Verantwortliche haben die Aufsichtsbehörde vorab einer Verarbeitung zu konsultieren, wenn aus einer Datenschutz-Folgenabschätzung gem. Art. 35 DS-GVO hervorgeht, dass die Verarbeitung ein hohes Risiko für Betroffene hätte, sofern keine Maßnahmen zu dessen Eindämmung getroffen werden. Liegt ein hohes Risiko für Betroffene vor und stellt die Pseudonymisierung die einzige Schutzmaßnahme dar, tritt die gesetzliche Pflicht zur Konsultation der zuständigen Aufsichtsbehörde ein.

2.1.9. Dokumentation und regelmäßige Evaluation des Prozesses, der erfolgten Abwägungen und der tatsächlich getroffenen Maßnahmen

Zu jedem Abschnitt des Kapitels 2.1 sind die getroffenen Maßnahmen sowie die relevanten Einflussfaktoren zu Festlegung einer angemessenen Pseudonymisierungsmethode (Abschnitt 2.1.2) zu dokumentieren. Soweit der Festlegung der getroffenen Maßnahmen eine Abwägung voranzustellen ist, sind derartige Abwägungen ebenfalls zu dokumentieren.

Die Dokumentation ist durch den FvfP sicherzustellen. Der FvfP kann aber auf Dokumentationen anderer Fachverantwortlicher, sowie auf Dokumentationen Dritter zurückgreifen. Hierbei ist sicherzustellen, dass Modifikationen der Dokumentation ausschließlich transparent erfolgen; insbesondere hinsichtlich der Aspekte „was“, „durch wen“ und „wann“.

2.1.9.1. Dokumentation von Prozessen und sonstiger getroffener Maßnahmen

Prozesse und getroffene Maßnahmen sind so zu dokumentieren, dass
1. der FvfP in der Lage ist,
   - den Prozess oder die Maßnahme hinsichtlich der Wirksamkeit zu bewerten;
   - die Implementierung der Prozesse oder der getroffenen Maßnahmen zu verifizieren;
   - die Einhaltung der Prozesse oder der getroffenen Maßnahmen zu evaluieren.
2. der FvfP und alle mit der Umsetzung betrauten Personen in der Lage sind,
   - den Prozess oder die Maßnahme zu verstehen und entsprechend der definierten Vorgaben umzusetzen.

[2] Hinweise zu Risikobestimmung finden sich z.B. im Kurzpapier Nr. 18 der Konferenz der unabhängigen Datenschutzbeauftragten des Bundes und der Länder oder im Working Paper 248 der Artikel-29-Datenschutzgruppe.

2.1.9.2. Dokumentation von Abwägungen

Abwägungen sind inklusive einer Begründung zu dokumentieren. Hierbei ist sicherzustellen, dass die im Rahmen der Abwägung getroffenen Schlussfolgerungen – z.B. Festlegung der geeigneten Pseudonymisierungsmethode oder eine angewandte Risikoklassifizierung – auch durch Dritte ohne Weiteres nachvollzogen werden können. Diese Abwägungen sind, insbesondere unter den Aspekten Stand der Technik sowie Zweckkonformität regelmäßig zu überprüfen und auch diese Überprüfungen sind zu dokumentieren. Re-ferenzen auf weitere Dokumentationen sind zulässig, soweit es sich um referenzierte Methoden handelt oder entsprechend dieses Abschnitts dokumentierte Abwägungsergebnisse und die Referenz den konkreten Titel, Speicher- oder Ablageort und Version des referenzierten Dokuments aufweist.

Erläuterung: Die nach diesem Abschnitt zu erstellende Dokumentation erfüllt mehrere Ziele. Die Dokumentation zwingt den Verantwortlichen bzw. den Auftragsverarbeiter, die Vorgaben dieses Codes systematisch zu bearbeiten. Soweit der FvfP auf Zuarbeiten anderer Fachverantwortlicher zurückgreift, verfügt der FvfP über eine auch für sich stets nachvollziehbare Informationsbasis. Weiterhin im Nachgang ermöglicht diese Dokumentation dem FvfP die ursprünglichen Annahmen regelmäßig zu überprüfen und bei Bedarf anzupassen. Eine solche Evaluation ist insofern erforderlich, als dass die Datenschutz-Grundverordnung eine Verarbeitung nach dem jeweiligen Stand der Technik verlangt. Mithin ist es wahrscheinlich, dass getroffene Maßnahmen oder auf der dokumentierten Informationsbasis erfolgte Abwägungen mit voranschreitendem technischen Fortschritt modifiziert werden müssen. Im Übrigen ermöglicht die Dokumentation sowohl dem FvfP sowie etwaigen Compliance-Abteilungen Konformitätsprüfungen durchzuführen.

2.2. Technische Fragen

2.2.1. Allgemeine Anforderungen an die Pseudonymisierung

Die technische Umsetzung findet ausschließlich in Rücksprache mit dem FvfP statt. Hierbei hat der FvfP bei der Auswahl und Bewertung der angemessenen Pseudonymisierungsmethode die technischen Fachverantwortlichen zu konsultieren. Ebenso haben die technischen Fachverantwortlichen den FvfP über geplante Änderungen der technischen Umsetzung zu konsultieren.

Zur Umsetzung einer Pseudonymisierung können unterschiedliche Verfahren eingesetzt werden. Beispielsweise kann eine Zuordnungstabelle verwendet werden, in der jedem Klartextdatum ein oder mehrere Pseudonyme zugeordnet werden. Alternativ können zur Pseudonymisierung verschiedene kryptographische Verfahren eingesetzt werden, die jeweils ein Klartextdatum in ein oder mehrere Pseudonyme überführen. Über den Zugriff auf die verwendeten kryptographischen Schlüssel und ggf. weitere Parameter kann hier die Umkehrbarkeit der Pseudonymisierung gesteuert/eingeschränkt werden.

Bei der Wahl der einzusetzenden Pseudonymisierung sind die Prüfschritte der Bestandsaufnahme (insbesondere Ziff. 2.1.2.1. bis 2.1.2.6.) zu durchlaufen.

2.2.2. Allgemeine Anforderungen an Identifikatoren (IDs)

Ungeachtet der weiteren Anforderungen ist als Pseudonym eine ID zu verwenden, die in sich keine Rückschlüsse auf die Eingabedaten oder die ggf. betroffene natürliche Person zulässt.

Anwendungsszenarien und Herausforderungen:

- Bei der Pseudonymisierung von Daten muss sichergestellt werden, dass die genutzte ID nicht re-identifiziert werden kann, wenn Einzelinformationen im Datensatz im Kontext mit anderen Daten betrachtet werden.

Erläuterung: Es wird die Postleitzahl als ID verwendet; weiterhin enthalten die Daten Einzelinformationen zum Geburtsdatum. Bei einer hinreichend kleinen Anzahl von Datensätzen ist eine Re-Identifikation der natürlichen Person durch Vergleich aller Datensätze mit identischem Geburtsdatum möglich.

- Wenn IDs auf Grundlage der Kombination von Einzelinformationen in den betrachteten Datensätzen generiert werden, ist sicherzustellen, dass bei einem direkten Vergleich der Ausgabedaten mit den Eingabedaten oder bei Kenntnis des verwendeten Schemas eine Re-Identifizierung nicht mit einfachen Mitteln zu erreichen ist. Dies kann z.B. durch Hinzuziehen eines geheimen Schlüssels („Salt“) bei der Berechnung der Pseudonyme erreicht werden.
- Es sind Verfahren zu bevorzugen, die keine Rückschlüsse auf die Sortierung der Daten zulassen oder die Sortierung der Daten vor Anwendung der Verfahren muss ausreichend zufällig erfolgen.

Erläuterung: Pseudonymisierte Daten könnten sich durch eine leicht nachzuvollziehende chronologische oder alphabetische Abfolge re-identifizieren lassen.

Bezüglich der verwendeten technischen Verfahren sind darüber hinaus die relevanten aktuellen technischen Richtlinien allgemeiner Art zu berücksichtigen, im Besonderen die relevanten Richtlinien des BSI („TR-02102 Kryptographische Verfahren: Empfehlungen und Schlüssellängen“), wenn z.B. Verfahren genutzt werden, die als Basis Hashfunktionen einsetzen. Für die Pseudonymisierung eingesetzte Transformationsverfahren müssen zudem – insbesondere bei langfristig verwendeten pseudonymisierten Daten – durch jeweils aktuelle Verfahren ausgetauscht werden, um ein Höchstmaß an Sicherheit zu gewährleisten.

2.2.3. Berechnungsverfahren

Die Wahl der spezifischen Pseudonymisierungsmethode muss auf Grundlage der Bestandsaufnahme erfolgen und mit dem FvfP abgestimmt werden; entsprechend unterliegt auch die technische Umsetzung der regelmäßigen Evaluation, vgl. 2.1.9.2.

Bei dem Einsatz von Berechnungsverfahren zur Bestimmung von Pseudonymen (insbesondere für pseudonyme Benutzer) ist sicherzustellen, dass diese folgende Eigenschaften haben:

1. Sie müssen auf nach dem aktuellen Stand der Technik sicheren kryptographischen Verfahren basieren.

Erläuterung: Software zur Bildung der Pseudonyme sollte verfügbare Krypto-Bibliotheken verwenden, anstatt die Algorithmen neu zu implementieren. Deswegen sind z.B. Open Source-Implementierungen sinnvoll.

2. Für den gegebenen Klartext-Raum (z.B. die Menge aller User-IDs oder Namen oder Telefonnummern) muss die Funktion Pseudonym = f(Klartext-ID) eindeutig sein, d.h. bei unterschiedlichen Klartext-Schlüsseln müssen unterschiedliche Pseudonyme resultieren, um Homonymfehler sicher zu vermeiden.

Erläuterung: Ein Homonymfehler entsteht, wenn bei Verkettbarkeit leistenden Pseudonymisierungsverfahren Identitätsdaten von unterschiedlichen Personen fälschlicherweise zu gleichen Pseudonymen führen.

3. Die Umkehrfunktion Klartext-ID = g(Pseudonym) darf nicht mit vertretbarem Aufwand berechenbar sein.

Erläuterung: Der zu Grunde gelegte vertretbare Aufwand sollte hierbei auch auf Basis der konkreten Begebenheiten ermittelt werden. Hierbei sollte insbesondere Berücksichtigung finden, welchen Wert die re-identifizierten Daten für unberechtigte Parteien haben. Hierzu kann auf die erfolgte Risikoanalyse zurückgegriffen werden. Diese Information ist insoweit für die Ermittlung des vertretbaren Aufwands wichtig, da diese Rückschlüsse auf die zu erwartenden technischen und fachlichen Ressourcen unberechtigter Parteien zulässt: Je höher der Wert der Daten, desto größer ist der aus Sicht unberechtigter Parteien vertretbare Aufwand.

4. Ähnliche, insbesondere aufeinanderfolgende Klartext-IDs dürfen nicht zu ähnlichen Pseudonymen führen, kleine Änderungen an Klartext-IDs müssen zu völlig unterschiedlichen Pseudonymen führen, um die Möglichkeit des „Erratens“ von Klartext-IDs zu erschweren.

5. Die Sicherheit der Pseudonymisierung darf nicht durch Geheimhalten des Algorithmus erreicht werden, sondern durch einen geheimen Schlüssel.

6. Aus der Kenntnis eines Paars (Klartext-ID/Pseudonym) darf nicht mit vertretbarem Aufwand auf den eingesetzten geheimen Schlüssel geschlossen werden können.

7. Aus den Punkten 1.-6. resultiert die Empfehlung, die Pseudonymisierung mithilfe einer kryptographischen Hash-Funktion oder eines symmetrischen Blockchiffreverfahrens durchzuführen, bei dem neben der Klartext-IDs ein geheimer, konstanter Schlüssel eingeht, dessen Entropie mindestens 100 Bit beträgt. Entropie bezeichnet ein Maß für die Unbestimmtheit einer Zeichenfolge (z.B. liefern zehn voneinander unabhängige Münzwürfe (Kopf/Zahl) zehn Bit Entropie). Bei Verwendung einer Hashfunktion ergibt sich die Mindestlänge des Hashwerts aus der Forderung zu in Ziff. 3.

3. Anwendungsbeispiele der Pseudonymisierung

3.1. Pseudonymisierung Magenta TV (DTAG)

3.1.1. Einleitung

Die Deutsche Telekom erzeugt anonyme Statistiken, die auf der Nutzung des Produktes Magenta TV beruhen. Dabei werden personenbezogene Daten zunächst pseudonymisiert, um diese in eine anonyme Statistik zu überführen. Für die Pseudonymisierung werden insbesondere bestimmte Nutzungsdaten, so genannte Events, verwendet, die mit einem Identifier (ID) versehen sind. Damit ist beispielsweise eine unterschiedliche Zählung möglich. D.h. es kann sowohl die Frage beantwortet werden, wie viele Haushalte oder wie viele Set-Top-Boxen zu einer bestimmten Zeit einen bestimmten Sender gesehen haben. Jeder Nutzer hat zu jederzeit die Möglichkeit, dieser Verarbeitung zu widersprechen (Opt-out).
In der anonymen Statistik sind diese IDs letztlich nicht mehr vorhanden, somit ist eine Rückführung von den reinen Zahlen auf die verschlüsselten IDs unmöglich.

3.1.2. Beschreibung der Verantwortlichkeiten

Verantwortlich für die personenbezogenen Daten, welche bei der Nutzung des Produktes Magenta TV erzeugt werden, ist die Telekom Deutschland GmbH. Die Pseudonymisierung wird von der T-Systems GmbH als IT-Dienstleister vorgenommen. Dabei wird die T-Systems von der Telekom Deutschland über einen Auftragsverarbeitungsvertrag eingebunden. Über eine weitere Legaleinheit der T-Systems, die Tel-IT, wird ein automatisiert erzeugter Schlüssel für die Pseudonymisierung bereitgestellt. Zudem ist sie an der Entwicklung und am Betrieb beteiligt.

Die Beauftragung der Pseudonymisierung wird von dem Segment „Privatkunden Deutschland“ vorgenommen. D.h. dieser Bereich beauftragt die Pseudonymisierung, nach Absprache mit dem Konzerndaten-schutz der Deutschen Telekom, bei dem IT-Dienstleister. Dieser Bereich ist auch zuständig für die Rechtskonformität des Pseudonymisierungs-Prozesses als solchem.

3.1.3. Kriterien für die Bestimmung der geeigneten Pseudonymisierungsmethode

Bei den zu pseudonymisierten Datenarten handelt es sich um Nutzungsdaten von Magenta TV, vgl. Art. 4 Nr. 1 DS-GVO. Zudem gibt es noch Metadaten, welche ebenfalls in die Pseudonymisierung einfließen. Diese werden zur Erstellung von Nutzungsprofilen pseudonymisiert, vgl. Art. 6 Abs. 1 lit. f) i.V.m. Art. 32 Abs. 1 lit. a) DS-GVO. Dabei handelt es sich pro Tag um mehr als 10 Millionen Datensätze, welche pseudonymisiert werden. Für die Pseudonymisierung werden Personen- und Geräte-Pseudonyme erstellt.

Datenfeld NAME	KENNUNG	Risiko-Klasse	Bemerkungen
Subscri-ber_ID	ACCOUNT_ID	1	Pseydonym Abonnent
Physical_ Device_ID	DEVICE_ID	2	Pseydonym Geräte-Kennung

Abb.: Datenfelder und Risikoklassen

3.1.4. Rechte- und Rollenkonzept sowie Schlüsselverwaltung

Die Berechtigungen sind sowohl durch Rollenbindung als auch technische Zweckbindung klar verteilt, welches im Organisations- und Berechtigungskonzept niedergelegt ist. Der Bereich der Telekom Deutschland GmbH (TDG), der verantwortlich ist für das Produkt Magenta TV, hat keinen Einfluss auf die Pseudonymisierung. Er hat lediglich Zugriff auf die erzeugten anonymen Statistiken, welche am Ende erzeugt werden. T-Systems nimmt die Pseudonymisierung vor, indem über den AcL (Acquisition Layer) die Nutzungsdaten automatisiert verschlüsselt werden.

Eine unabhängige technische Instanz (Tel-IT) liefert den Schlüssel. Auf dieses System haben nur die Tel-IT und die Administratoren der T-Systems Zugriff. Das für die Pseudonymisierung benötigte Kryptomaterial (Schlüssel/Salts) ist separat gekapselt, in einem so genannten Trust Center (Tel-IT). Bei der Konfiguration hat der Mitarbeiter keine Möglichkeit, Kenntnis davon zu erlangen. Auf das Kryptomaterial haben nur technische Nutzer sowie ein kleiner gesonderter Personenkreis (3-4 Personen) Zugriff. Diese haben aller-dings keine Adminrechte. Hier liegt die organisatorische Trennung vor.

Zudem verzichtet die TDG in einer zusätzlichen Vereinbarung auf die Weisungshoheit in Bezug auf das Kryptomaterial, die sie gemäß der Auftragsdatenverarbeitung hätte, D.h. TDG darf diese Informationen nicht anfordern. Tel-IT darf sie nicht herausgeben, auch nicht an Dritte. Erst wenn die Pseudonymisierung abgeschlossen ist, werden die Daten vom AcL an die BDMP (Big Data Management Plattform) übermittelt und stehen dort der TDG für Analysen zur Verfügung. Auf der BDMP werden die pseudonymisierten Nut-zungsprofile aggregiert. Ein Zugriff auf die Information im AcL sowie die technische Instanz ist ausgeschlossen.

3.1.5. Datenerzeugung

Beim Benutzen einer Magenta TV Set-Top-Box (STB) – d.h. beim Drücken der Fernbedienung durch den Benutzer werden unterschiedliche Events generiert, je nachdem welche Tasten gedrückt wurden und in welchem Kontext der Benutzer sich befindet. Diese Events der STB stellen die Basis der Auswertungen dar.
Beispiele für diese Events sind z.B. Ein-/Ausschaltvorgänge, Kanalumschaltungen, Informationen zu den gesehenen Sendern oder Informationen zu Aktivitäten rund um das Aufnehmen bzw. Anschauen von Aufnahmen.

Diese Event-Datensätze enthalten z.B. Information über die Set-Top-Box (=DeviceID), Datum/Uhrzeit sowie weitere spezifische Datenfelder. Die personenbeziehbaren Informationen dieser Events werden mittels einer auf dem AES128[3] basierenden formaterhaltenden Chiffre verschlüsselt.

[3] Advanced Encryption Standard mit einer Schlüssellänge von 128 Bit.

3.1.6. Pseudonymisierung

Das zu Grunde liegende Pseudonymisierungsverfahren führt zu verkettbaren, aber nicht-aufdeckbaren Pseudonymen. Diese werden unter Verwendung so genannter deterministisch, kryptografisch starker Chiffren erstellt. Da deterministische Verfahren gleiche Klartexte auf gleiche Ergebniswerte (Pseudonyme) abbilden, ist die Verkettbarkeit sichergestellt. Durch die sichere Verwaltung des Schlüsselmaterials sowie organisatorische Trennung des Zugriffs auf die Schlüssel ist die unzulässige Umkehrung der Pseudonymi-sierung also die Aufdeckung des Klardatums ausgeschlossen.
Die für die AccountID (ID für den Kunden) und die DeviceID (ID für die jeweilige Set-Top-Box) gebildeten Pseudonyme werden für die weiteren Auswertungen verwendet. In den zur Auswertung notwendigen Event-Dateien und den Referenzen ACCOUNT_PS und DEVICE_PS sind keine Attribute enthalten, welche direkt personenbezogene Daten enthalten. Diese Referenzen (ACCOUNT_PS und DEVICE_PS) sind die Personen- und Geräte-Pseudonyme.
Die Pseudonyme werden verwendet, um die Nutzungsinformation von Magenta TV zu erfassen, um damit anonyme Statistiken zu generieren. Hierbei ist es wichtig, erkennen zu können, welches Ereignis vom selben Gerät oder Nutzer erfolgt. Durch die Pseudonymisierung wird sichergestellt, dass Mitarbeiter keine Rückschlüsse auf die tatsächlichen Geräte oder Nutzer ziehen können. Die im Anschluss daraus generierten Statistiken werden vollständig von den pseudonymisierten Kennern befreit und sind somit anonym.

3.2. Pseudonymisierung bei Optimierung der Online-Plattform-Werbung

3.2.1. Einleitung

Werbung über Online-Plattformen wie etwa Social-Media, E-Commerce-Shop oder Online-Publisher an eine gewünschte Zielgruppe zu richten, ermöglicht die Minimierung des Streuverlusts der Werbung. Gleichzeitig erspart zielgerichtete Werbung Plattform-Nutzern unnötige Irritationen durch irrelevante Werbevideos. So hilft die Nutzung der handelsüblichen Verbraucher-Informationen wie soziodemographische oder Lifestyledaten, relevantes Werbepublikum zu erreichen. Acxiom lizenziert mit Selektionskriterien gebildete Zielgruppen und nutzt dabei mehrfache Pseudonymisierung, damit die Daten zum einen verkettbar sind zum Zweck der Darstellung individualisierter Online-Werbung, zum anderen um Betroffene vor einer direkte Identifizierung zu schützen.

3.2.2. Vorbereitung: Erstellung einer Pseudonym-zu-Pseudonym-Referenztabelle mit dem Plattform-Partner

Um die gewünschten Zielgruppen online auf einer Plattform zu erreichen, ist ein zweistufiger Prozess erforderlich. Zunächst und unabhängig von einem konkreten Kundenauftrag findet ein Datenabgleich der Adressbestände von Acxiom und dem Plattform-Betreiber statt. Im ersten Schritt wird dabei die Adressda-tenbank auf ein Acxiom-eigenes proprietäres Privacy Enhancement Tool (PET) geladen. Dort erhält jeder Datensatz der Adress-

datenbank einen pseudonymen Personenschlüssel. Dieser Personenschlüssel wird nochmals mit einem Salt verhasht. Darüber hinaus pseudonymisiert Acxiom die Klardaten der Adress-Datenbank, indem sie diese verhasht. So entsteht eine Datei mit zwei Feldern: der verhashte Personenschlüssel und die verhashten Adressdaten (Matchdatei-Acxiom). Der Plattform-Betreiber auf der anderen Seite pseudonymisiert seine Nutzerdaten auf analoge Weise und speichert die verhashten Nutzer-Kontaktdaten mit der Plattform-eigenen Nutzer-ID in einer Datei (Matchdatei-Plattform). Nach dem Ab-gleich der beiden Matchdateien auf Basis der Pseudonyme bzw. den Hashwerten wird eine Referenz zwischen der Plattform-Nutzer-ID und dem verhashten Acxiom-Personenschlüssel erstellt. Der Plattform-Betreiber speichert lediglich die Zuordnung der Plattform-Nutzer-ID zum verhashten Acxiom-Personenschlüssel. Alle anderen Informationen werden direkt im Anschluss an den Abgleich gelöscht.

3.2.3. Zielgruppenselektion

Für die Selektion der relevanten Zielgruppen auf einer Plattform wird bei Acxiom ein eigenes Datenprodukt erstellt, das die Pseudonyme (ein mit einem bestimmten Salt verhashten Acxiom-Personenschlüssel) als Schlüsselvariable enthält, jedoch keine Namen oder Anschriften.

Zielgruppen können anhand soziodemographischer Daten, berechneter Affinitäten für bestimmte Produkte oder Dienstleistungen, aber auch auf Basis rein geographischer Information (z.B. Werbung für Highspeed-Internet nur in Regionen, in denen dieses auch verfügbar ist) selektiert werden. Acxiom verfügt über eine breite Palette an mikrogeographischen Merkmalen, die anhand amtlicher Daten, Umfragen und Marktforschungsstudien etc. auf feinräumiger Nachbarschafts-Ebene berechnet werden (d.h. alle Haushalte einer geographischen Zelle bzw. Nachbarschaft bekommen alle die gleichen Werte zugeordnet). So wird z.B. die Annahme „besitzt eine Katze“ allen Haushalten dieser mikrogeographischen Zelle, unabhängig von der individuellen Situation der verschiedenen Familien in der Nachbarschaft, die stets mind. 4 Haushalte umfassen muss[4], zugeordnet. Durch das Nutzen dieser Merkmale wird die Identifizierung bzw. Aufdeckbarkeit einer natürlichen Person mittels dieser pseudonymen Datensätze verhindert.

Ergebnis einer Zielgruppenselektion (z.B. „besitzt eine Katze“ und „wohnt in einer Wohnung“) ist stets eine Liste aus verhashten Acxiom-Personenschlüsseln. Diese wird von Acxiom in den Acxiom-Werbeaccount beim Plattform-Betreiber hochgeladen und kann von dort mit dem Werbetreibenden bzw. seiner Agentur geteilt werden, so dass diese die Zielgruppe nutzen können.

3.2.4. Schalten von Werbung

Anhand der im Datenabgleich erstellen Referenz zwischen der Plattform-Nutzer-ID und dem verhashten Acxiom-Personenschlüssel schaltet der Plattform-Betreiber die Werbeeinblendung in den Accounts der hochgeladenen und geteilten Zielgruppe.

4 Gemäß den Empfehlungen des 3. Geo-Fortschrittsberichts der Bundesregierung aus Oktober 2012.

3.2.5. Technische und organisatorische Maßnahmen

Der Plattform-Betreiber hat sich gegenüber Acxiom vertraglich verpflichtet, die Referenzdaten zwischen der Plattform-Nutzer-ID und verhashtem Acxiom-Personenschlüssel separat und physisch getrennt von seinem CRM-System zu halten. Die Ausspielung der Werbung erfolgt über ein separates Werbeauslieferungssystem. Durch die entsprechende vertragliche Verpflichtung sowie die physisch von seinem eigenen Nutzerdatenbestand getrennte Verarbeitung der Daten ist sichergestellt, dass auch für den Plattform-Betreiber keine Möglichkeit besteht, auf Basis dieser pseudonymen IDs eine Person zu identifizieren.

Bei Acxiom ist die Zugangsberechtigung zum Salt, der zum Verhashen des Personenschlüssels verwendet wird, nur wenigen ausgewählten Mitarbeiter erteilt. Für die die Zielgruppen selektierenden Acxiom-Mitarbeiter sind die ausgewählten ID-Nummern nicht aufdeckbar, da sie auf Grund des oben beschriebenen Prozesses keine Möglichkeit haben, die verhashten Acxiom-Personenschlüssel einer Person zuzuordnen.

Darüber hinaus sind die zweifach pseudonymisierten und verhashten Acxiom-Personenschlüssel sowohl für die Werbetreibenden, die die Zielgruppen von Acxiom lizensieren, als auch für deren Agentur wie auch für jeden anderen Dritten anonym, da sie keine Möglichkeit haben, hieraus eine Person zu identifizieren oder sie einem Individuum zuzuordnen. Zudem haben die Werbetreibenden schon keinerlei Zugriff auf die verhashten Acxiom-Personenschlüssel in der selektierten Zielgruppe, denn die Auswahl der Zielgruppen sowie das Hochladen auf die Plattform findet ausschließlich bei Acxiom stattfindet. Ein Einblick in die hochgeladene Zielgruppe ist dem Werbetreibenden schon technisch nicht möglich.

Anlässlich des Digital-Gipfels 2019 hat es sich die Fokusgruppe Datenschutz der Plattform 9 „Sicherheit, Schutz und Vertrauen für Gesellschaft und Wirtschaft“ zur Aufgabe gemacht, einen Entwurf für einen Code of Conduct für die Pseudonymisierung personenbezogener Daten zu erarbeiten. In Zeiten enormer Datenmengen, mit denen Anwendungen der künstlichen Intelligenz oder des maschinellen Lernens gefüttert werden können, kann die Pseudonymisierung einen wichtigen Beitrag leisten, einen Ausgleich zwischen technologischem Fortschritt und den Persönlichkeitsrechten von Nutzerinnen und Nutzern zu schaffen.