# ANONYMISIERUNG: METHODEN UND ZULÄSSIGKEIT

Michael Sonntag

Assoz.-Professor, Johannes Kepler Universität Linz, Institut für Netzwerke und Sicherheit
Altenbergerstr. 69, 4040 Linz, AT
michael.sonntag@ins.jku.at; http://www.ins.jku.at/

***Schlagworte:*** ***Anonymisierung, Methoden, Zulässigkeit***

***Abstract:*** *Die Bedeutung von Anonymisierung im Internet steigt, da immer mehr Daten über Personen gespeichert werden, wobei oft unklar ist, wo diese sind, wer darauf Zugriff hat, und wer in Zukunft darauf, u.A. auch illegal z.B. durch Hacking, Zugriff erlangen wird. Dieser Beitrag stellt dar, mittels welcher grundlegenden Techniken Anonymisierung erfolgt und untersucht ihre Zulässigkeit, sowohl für Betreiber von Anonymisierungsdiensten wie auch im Hinblick auf ihren Einsatz durch Nutzer.*

## 1. Einleitung

Wie die Ereignisse um das Seitensprung-Portal Ashley-Madison gezeigt haben, ist Anonymität im Internet ein Aspekt, den man oft vergisst, bis es zu spät ist. Diese Website zur Vermittlung von Seitensprüngen wurde gehackt. Nachdem Forderungen, den Dienst zu schließen, nicht erfüllt wurden, erfolgte die Veröffentlichung aller Kundendaten im Internet. Dies ist gleichzeitig ein Beispiel wider das Mantra gegen Datenschutz, «ich habe nichts zu verbergen». Denn obwohl ein Seitensprung legal ist (d.h. nicht mit Gerichts- oder Verwaltungsstrafe bedroht), möchten doch viele NutzerInnen eines solchen Portals nicht, dass ihre (versuchte) Aktivität allgemein oder auch nur dem Partner bekannt wird, selbst wenn man derzeit/damals nicht in einer Beziehung lebt/e.

Was wäre erforderlich, um eine derartige Website anonym zu nutzen? Offensichtlich wäre auf identifizierbare Fotos zu verzichten. Praktisch dürfte es für sehr viele Personen unmöglich sein, gezielt nach einem Bild als «Suchbegriff» zu suchen (siehe jedoch die Personenerkennungs-Algorithmen z.B. von Facebook!). Weiters wäre eine nicht rückverfolgbare E-Mail-Adresse zu verwenden. Diese E-Mail-Adresse dürfte ebenso nicht für andere Angelegenheiten verwendet werden. Weiters sollten im Portal keine zu genauen Angaben gemacht werden, denn z.B. alleine über Wohnort und tagesgenaues Geburtsdatum ist eine Identifikation leicht möglich. Surft man die Site vom Arbeitsplatz aus an, so verbleiben Spuren in Logdateien (hier anscheinend nicht vorhanden/nicht veröffentlicht). Anonymität ist daher mehrfach wichtig: Kommunikationsendpunkte (E-Mail-Adresse), Daten (Bilder, Wohnort, Zeiten und Treffpunkte) und Verhalten (IP-Adressen bei Kommunikation). Kann oder will man der Website nicht vertrauen, sind alle diese Elemente zu anonymisieren, um eine Identifizierung zu verhindern.

## 2. Methoden von Anonymisierung

Bei Anonymisierung sind mehrere Aspekte wichtig: Es können Daten anonymisiert werden (z.B. der Inhalt einer Datenbank, DB, auf einem Server), aber auch Kommunikation, wobei diese wiederum zwei Elemente enthält: Den Kommunikationsinhalt (z.B. den Mailtext) sowie den Kommunikationsvorgang selbst bzw. dessen Endpunkte (wie Absender- und Empfängeradresse). Für alle Aspekte sind unterschiedliche Techniken erforderlich, deren Grundprinzipien hier dargestellt werden.

## 2.1. Anonymisierung von Daten

Hierbei handelt es sich typischerweise um Einträge in Datenbanken, deren Inhalte auf Personen zurückführbar sind. Folgende Ansätze zur Anonymisierung sind bekannt[1,2]:

- Entfernen identifizierender Merkmale[3]: Diese unfehlbare Methode führt allerdings meist dazu, dass die Daten nutzlos werden. Weiters kann aus mehreren «harmlosen» Datenfeldern oft Re-Identifikation erfolgen, wenn Datenteile aus anderen Quellen bekannt sind. Eine verminderte Variante ist Partitionierung, bei welcher identifizierende Daten separat gespeichert werden. Während der erste Ansatz anonymisiert, ist Partitionierung eine techn. Sicherheitsmaßnahme ohne Änderung der rechtl. Qualifizierung (weiter direkt personenbezogen).

- Pseudonyme[4]: Identifizierende Informationen werden durch geheime Daten ersetzt. Dies ähnelt dem vorigen Punkt (Partitionierung) und besitzt ähnliche Schwächen. Weiters ist zu berücksichtigen, dass z.B. das Ersetzen von Daten durch Hashwerte es ermöglicht, die Daten einer bestimmten Person, von welcher die gehashten Werte bekannt sind, abzurufen. Ein Datensatz kann daher keiner Person zugeordnet werden, aber für bekannte Personen können die Daten – sofern vorhanden – trivial festgestellt werden.

- Randomisieren[5]: Einzelne Datenfelder werden zwischen Datensätzen vertauscht, z.B. die Adresse wird einer anderen Person zugeordnet. Wie im vorigen Beispiel ersichtlich, ist dies nicht unbedingt eine Anonymitätsgarantie. Für statistische Auswertungen einzelner Felder ist dies gut geeignet, doch wäre hier die Erstellung der Statistik und anschließende Löschung aller Grunddaten ausreichend und sicher.

- Unschärfe einfügen[6]: Datenfelder werden leicht verändert, wobei darauf geachtet wird, statistische Merkmale beizubehalten. So kann z.B. das Geburtsdatum um mehrere Jahre hinauf oder hinunter gesetzt werden. Dies ermöglicht weiterhin Auswertungen mehrerer Felder (mit leicht erhöhter Ungenauigkeit) sowie statistische Auswertungen einzelner Felder. Doch auch eine Re-Identifizierung ist oft weiterhin durchführbar.

- Eine Variante des vorigen Punktes ist die Generalisierung[7]: Daten werden nicht zufällig verändert, sondern auf größere Kategorien abgebildet, z.B. die Wohnadresse auf den Wohnort. Damit gehen Daten verloren, aber eine Re-Identifizierung ist durch Kombination mehrerer Felder vielfach weiterhin möglich.

- Aggregation[8]: Zusammenfassung mehrerer Elemente zu Gruppen. Sind die (alle!) Gruppen groß genug, so ist es nur mehr schwer möglich, auf Einzelpersonen zu schließen. Statistische Aussagen über die

[1] Siehe dazu mit Beispielen: Art. 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques, 10. April 2014, 0829/14/EN, WP 216; sowie Cormode/Srivastava, Anonymized Data: Generation, Models, Usage, Proceedings of the 2009 ACM SIGMOD International Conference on Management of data, ACM 2009.

[2] Mit noch stärkerer Unterteilung: Nelson, Practical Implications of Sharing Data: A Primer on Data Privacy, Anonymization, and De-Identification, SAS Global Forum 2015.

[3] Sweeney, k-anonymity: a model for protecting privacy, International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, 10 (5), 2002, 557–570.

[4] Kuenning/Miller, Anonymization Techniques for URLs and Filenames, Technical Report UCSC-CRL-03-05, 2003, http://www.ssrc.ucsc.edu/Papers/crltr-03-05.pdf (alle Websites zuletzt besucht am 7. Januar 2016).

[5] Zhang/Koudas/Srivastava/Yu, Aggregate query answering on anonymized tables, IEEE International Conference on Data Engineering, IEEE 2007.

[6] Aggarwal, On unifying privacy and uncertain data models, IEEE International Conf. on Data Engineering, IEEE 2008.

[7] Klösgen, Anonymization techniques for Knowledge Discovery in Databases, Proceedings of First International Conference on Knowledge Discovery & Data Mining, KDD-95, AAAI 1995, 186–191.

[8] Torra, Constrained microaggregation: Adding constraints for data editing, Transactions on Data Privacy 1(2), 86–104 (2008).

Gruppen sind hingegen weiterhin möglich. Zusätzlich ergibt sich das Problem, dass die Gruppierung immer gleich bleiben muss, da sonst die Anonymität geschwächt wird. Um wirklich gute Anonymisierung zu erreichen, sind besondere Strategien erforderlich (alle Gruppen müssen annähernd gleich groß sein, die Verteilung der Werte innerhalb der Gruppe ist zu berücksichtigen etc.).

Besonders problematisch im Hinblick auf Anonymisierung ist, dass sie zwar relativ gute Ergebnisse (nach beachtlichem Aufwand!) erlauben, wenn es um einmalige Auskunft/Veröffentlichung geht. «Lebt» die DB und kommen Daten hinzu bzw. werden entfernt und erfolgt eine Anonymisierung (und z.B. anschließend eine Veröffentlichung) hingegen wiederholt, so stellen sich besondere Probleme: Die frühere Anonymisierung erforderte aufgrund der zwischenzeitlichen Änderungen u.U. andere Parameter (Unschärfe, Gruppierung[9] etc.) als die spätere. Sind jedoch beide anonymisierten Ergebnisse verfügbar, lassen sich daraus leichter die Ursprungsdaten wiederherstellen. Gleiches gilt, wenn aus einer nicht-anonymen DB mehrere unterschiedlich anonymisierte Extrakte generiert werden: Zusammen können diese eine Rekonstruktion der Ursprungsdaten erlauben.

Rechtlich ist dies besonders bedeutsam, da zumindest derzeit immer nur auf den konkreten Einzelvorgang abgestellt wird: Frühere Veröffentlichungen werden u.U. durch das «allgemeine Wissen» miteinbezogen, doch Zukünftige finden keine Beachtung. Eine isolierte Betrachtung jedes Anonymisierungsvorganges ist daher aus Sicht eines effektiven Datenschutzes abzulehnen.

## 2.2. Anonymisierung von Kommunikationsinhalten

Die Anonymisierung von Kommunikationsinhalten ähnelt der Anonymisierung von Daten. Doch im Gegensatz zu dort handelt es sich um Daten einer Einzelperson. Es ist daher nicht möglich, diese Person unter vielen zu «verstecken», sondern es sind tatsächlich alle identifizierenden Informationen zu entfernen, da ansonsten Anonymität nicht erreichbar ist. Als Beispiel kann eine E-Mail dienen: Der gesamte Inhalt (inkl. der Header) muss bearbeitet werden um sicherzustellen, dass daraus kein Rückschluss auf den Absender mehr möglich ist. Praktisch stellen sich Probleme, derartige Informationen automatisch zu identifizieren bzw. Benutzer darüber zu unterrichten, welche Angaben zu unterlassen oder welche Vorkehrungen zu treffen sind.

Am sichersten ist hier die Verwendung reiner ASCII-Text-Kommunikation mit besonderem Augenmerk auf darin verwendete Informationen[10]. Eine vollständige End-to-end-Verschlüsselung[11] schützt hingegen nur auf dem Übertragungsweg[12], aber nicht am Ziel.

## 2.3. Anonymisierung von Kommunikationsvorgängen

Abgesehen vom Kommunikationsinhalt im weitesten Sinne kann der Vorgang selbst ebenso zu Identifikation führen. Denn selbst wenn der Inhalt vollständig verschlüsselt ist, besitzt eine Übertragung über das Internet zwingend eine korrekte Ziel-IP-Adresse. Sind die Daten nicht extrem kurz bzw. sind Datenverluste relevant,

[9] Werden z.B. aus einer Gruppe von 100 Personen 99 entfernt, so ist die verbleibende Person nicht mehr anonym und muss einer anderen Gruppe zugeschlagen werden. Ist diese neue Gruppe allerdings ansonsten vollständig unverändert, ist dadurch nichts gewonnen und auch diese Gruppe muss umgestaltet werden.

[10] Auch daraus kann eine Identifikation möglich sein, z.B. wenn diese Kombination von Daten nur einer einzigen Person bekannt war. Dies ist das gleiche Prinzip wie bei individuellen Wasserzeichen: Die Rückführbarkeit auf eine Einzelperson, da exakt diese Daten ausschließlich dieser einen Person zur Verfügung standen.

[11] Bei E-Mails bedeutet dies nicht nur die Verschlüsselung des Inhalts sondern auch des Übertragungsvorgangs, da sonst z.B. Header nicht verschlüsselt übermittelt werden.

[12] Und dort nur eingeschränkt: Wird auf eine Identifizierung des Partners verzichtet (bei Anonymisierung Grundvoraussetzung!), so sind Man-in-the-Middle-Angriffe immer möglich. Die einzige Abhilfe ist eine Überprüfung der «Identität» der Gegenstelle über ein z.B. bei einem persönlichen Treffen ausgetauschtes Geheimnis.

muss auch die Quell-IP-Adresse korrekt sein. Damit ist z.B. über ISP eine Identifikation des Endgerätebesitzers möglich. Daher sind diese Daten ebenfalls zu anonymisieren. Gleiches gilt für Telefonnummern, doch sind entsprechende Anonymisierungsdienste nicht bekannt[13]. Bei Anonymisierung derartiger paketvermittelter Daten wird grundsätzlich ein zusätzlicher Rechner dazwischengeschaltet, welcher Daten entgegennimmt und weiterleitet. Damit wird die ursprüngliche Quell-IP-Adresse durch seine eigene ersetzt. Gleichzeitig wird die «echte» Ziel-Adresse den empfangenen Daten entnommen und anstatt seiner eigenen Adresse (=Ziel des ersten Paketes) eingesetzt. Gefährlich ist, dass bei Beobachtung von Ein- und Ausgang eine Zuordnung der Pakete möglich und daher zu verhindern ist. Hierzu werden eingesetzt:

- Verzögerungen: Daten werden nicht sofort weitergesendet, sondern erst nach einer variablen Verzögerung. Damit wird eine zeitliche Analyse «Eingang zu Ausgang» erschwert. Dies ist z.B. bei E-Mails leicht möglich, schwerer beim Websurfen und fast unmöglich bei Chat/Internet-Telefonie (Zeitverzögerungen = Beeinträchtigung der Kommunikation).
- Auffüllen: Kurze Pakete werden aufgefüllt, damit aus der Länge der Daten keine Zuordnung möglich wird. Eine Verkürzung ist ebenso möglich, indem ein Paket in mehrere kleinere Teile aufgespalten wird. Gegebenenfalls wird bei fehlendem Datenverkehr solcher simuliert, um «Lücken» zu vermeiden.
- Vertauschung: Die Paket-Reihenfolge wird vertauscht und diese erst am Ende wieder korrekt gereiht. Dies ist nur relevant, wenn Auffüllen/Aufspalten nicht immer und in gleich große Pakete erfolgt oder nicht verschlüsselt wird. Praktisch daher kaum existent.
- Mischen: Kommunikation einer Vielzahl an Personen wird über denselben Server vermittelt. Damit wird es schwieriger, eine Einzelperson nachzuverfolgen, da ein ausgehendes Paket zu einer Vielzahl eingehender Paketen gehören könnte. Je mehr Nutzer ein Server pro Zeiteinheit bedient[14], desto besser die Anonymisierung.
- Unterschiedliche Wege: Pakete einer Verbindung können unterschiedliche Server zur Weiterleitung verwenden. Bei Verschlüsselung ist dies nicht mehr relevant. Bedeutung besitzt es hingegen bei mehreren Verbindungen: Eine Person X kommuniziert nicht nur mit A, sondern auch mit B und C, woraus sich Schlüsse über sie ziehen lassen.
- Mehrere Stationen: Um eine Identifikation zu erschweren, werden mehrere Server hintereinander geschaltet, sodass zur Identifikation eine Korrelation Ein- zu Ausgabe bei allen erfolgen müsste. Alternativ und unabhängig von der Stationsanzahl ist es möglich, sowohl an Quelle (z.B. ISP des Benutzers) als auch Ziel (ISP des Servers) den Datenverkehr abzugreifen und zu korrelieren. Dies bedeutet, dass Personen mit Abhörmöglichkeiten eines großen Teils des Internets eine De-Anonymisierung leichter durchführen können.

Tor[15] zur Anonymisierung von Internetkommunikation setzt z.B. folgende Techniken ein: Auffüllen, Vertauschung, Mischen sowie mehrere Stationen. Unterschiedliche Wege gelten nur für mehrere Verbindungen, während Verzögerungen fehlen. Daraus ergibt sich gleichzeitig, dass Tor gegenüber statistischen Analysen angreifbar ist, sofern die Kommunikation an beiden Enden abgehört werden kann (zeitliche Korrelation der Pakete zwischen Nutzer und Tor-Entry-Node sowie Tor-Exit-Node und dem, z.B. Web-, Server). Fingerprinting-Angriffe sind zwar möglich, aber in der Praxis aufgrund der Gegenmaßnahmen (insb. durch die Paketaufteilung

[13] Als Ansatz siehe Uber, Introducing Phone Anonymization, http://newsroom.uber.com/sa/2015/04/phoneanonymization/, bei der allerdings lediglich die beiden Endpunkte die Telefonnummer des Partners nicht kennen, Uber jedoch beide.

[14] Beispiel: Wird der Server von einem einzigen Nutzer verwendet, ist jegliche zeitliche Verzögerung irrelevant. Je mehr Benutzer ihn daher gleichzeitig verwenden, desto kürzer können die Verzögerungen werden.

[15] https://www.torproject.org/.

i.V.m. Auffüllen sowie Vertauschung auf der Anwendungs- und nicht nur der Paket-Schicht) deutlich schwerer als vermutet[16].

E-Mail-Anonymisierungsdienste hingegen setzen sehr oft Verzögerung ein, gerade um derartige Angriffe zu verhindern. Diese sind daher «besser», besitzen aber oft Probleme im Hinblick auf Antworten bzw. richterliche Abfragen (oft nur eine Station[17], daher mittels Logfiles, sofern vorhanden oder verpflichtend geführt, «leicht» rückführbar).

## 3. Grenzen der Anonymisierung

Anonymisierung findet ihre Grenzen in der Nutzbarkeit: Wird zu viel «weggeschnitten», bleibt u.U. nichts Interessantes mehr übrig. Dies ist unvermeidbar und wird daher nicht weiter thematisiert.

Eine weitere Limitierung besteht typischerweise darin, dass das eine Bestätigung der Autorenschaft bzw. Identifizierung von Kommunikation leicht möglich ist, wenn das verwendete Endgerät untersucht wird: Der lokale Mail-Account enthält u.U. eine Kopie der anonym versendeten E-Mail, der Browser-Cache beinhaltet abgerufene Webseiten sowie den Verlauf etc. Auch Tools zur Anonymisierung können zumindest einen Hinweis liefern und als Element der Bestätigung dienen. Zur Sicherstellung der Anonymität im Sinne von «deniability» (die Möglichkeit, alles Abstreiten zu können), ist daher auch das Endgerät entsprechend abzusichern, z.B. durch Vollverschlüsselung, Hidden Volumes, Verwendung von Live-DVDs o.Ä.

Aufgrund der Komplexität ist ein weiteres praktisch auftretendes Problem, dass kein Fehler erlaubt ist: Einmal nicht anonym eine einzige E-Mail verschickt, und schon ist man identifiziert; ein Plugin/Browser-Toolbar/… geladen und die Anonymität geht verloren. Praktisch existieren viele «Fußangeln», bei denen leicht irgendwann einmal eine davon übersehen wird. Dies ist u.A. der Grund, warum z.B. das Tor-Netzwerk (Anonymisierung des Web-Browsens) einen kompletten Browser mit spezieller Konfiguration und Anpassungen zur Verwendung bereitstellt und empfiehlt.

Vorteilhaft ist hingegen, dass Anonymisierung ggf. auf bestimmte Angreifer zugeschnitten werden kann: Befürchtet man Spionieren durch andere MitarbeiterInnen, so reichen schon einfache Techniken aus, da diesen nur begrenzte Ressourcen und weitere Daten zur Verfügung stehen. Generell sollte jedoch eher darauf verzichtet werden, absichtlich die Anonymität zu schwächen – was einmal anonym ist bleibt es auch, wo jedoch eine Zuordnung möglich ist oder wird, ist die Anonymität endgültig und für immer beseitigt.

## 4. Legalität von Anonymisierung

Die Legalität von Anonymisierung kann in verschiedenen Gestalten untersucht werden, u.A. hinsichtlich der Verwendung entsprechender Tools oder in Bezug auf das Anonymisieren «fremder» Daten, um diese anschließend speichern, bearbeiten usf. zu dürfen.

### 4.1. Was sind «anonyme Daten»?

Anonyme Daten werden im Datenschutzgesetz (DSG) nicht definiert. Es lassen sich jedoch einige Punkte hierfür angeben: Es darf kein Personenbezug vorliegen, da es sich sonst um direkt personenbezogene Daten handelt. Eine Zuordnung darf jedoch auch für Dritte nicht möglich sein (indirekter Personenbezug). Es muss sich daher um Daten handeln, bei denen Niemand sie einer bestimmten oder bestimmbaren Person (mehr) zuordnen kann. Weiters muss dies permanent und unwiderruflich gelten: Kann die Anonymisierung rückgängig

[16] Juarez/Afroz/Acar/Diaz/Greenstadt, A Critical Evaluation of Website Fingerprinting Attacks, https://securewww.esat.kuleuven.be/cosic/publications/article-2456.pdf.

[17] Beispiel: http://anonymouse.org/anonemail.html.

gemacht werden, z.B. über die Speicherungsreihenfolge der Datensätze, so liegen keine anonymen Daten vor. Hierbei ist allerdings ErwG. 26 der DS-RL[18] zu berücksichtigen, wonach diese Unumkehrbarkeit nicht absolut gilt, sondern nur im Hinblick auf Mittel, die vernünftigerweise eingesetzt werden können[19]. Ebenso soll besonderes Spezialwissen irrelevant sein. Dies bedeutet, dass z.B. eine Person aus «anonymen» Datensätzen denjenigen einer bestimmten Person herauspicken kann, über die er/sie viel weiß, alle anderen jedoch nicht zuordenbar bleiben[20]. Gerade für diese eine betroffene Person kann dies besonders nachteilig sein, da es sich offensichtlich um jemand «Bekannten» handelt, der zusätzliche Informationen erhält, während eine «globale» De-Anonymisierung durch Dritte ohne besonderen Bezug zu Einzelpersonen eher irrelevant wäre. Dies ist daher m.M. nach einschränkend auszulegen, insb. da in Österreich die Kategorie der indirekt personenbezogenen Daten existiert (auch wenn bei diesen nur auf rechtlich zulässige Mittel abgestellt wird und nicht auf praktische Möglichkeiten).

Für eine enge Auslegung von «Anonymität» spricht § 15 OGH-Gesetz[21]: Laut Abs. 1 ist eine Veröffentlichung des Volltexts einer Entscheidung im RIS zu unterlassen, wenn ansonsten die Anonymität der Betroffenen nicht sichergestellt ist. Da allerdings laut Abs. 4 ohnehin alle Namen, Anschriften und erforderlichenfalls sonstige Orts- und Gebietsbezeichnungen, die Rückschlüsse auf die betreffende Rechtssache zulassen, zu löschen bzw. ersetzen sind, reicht dies offensichtlich zur Anonymisierung nicht aus. Der Gesetzgeber geht daher davon aus, dass die Namenslöschung einen gewisser Schutz gewährleistet, dies aber keine vollständige Anonymisierung ist, da z.B. durch Zusatzwissen, welches u.U. weit verbreitet sein kann, eine Zuordnung weiterhin möglich bleibt. Dies spräche dafür, dass auch Spezialwissen dazu führt, dass keine anonymen Daten mehr vorliegen.

Bei Anonymisierung sollte daher in zwei Fälle unterschieden werden: Absolut anonymisierte Daten, die von niemandem zurückgeführt werden können, und relativ anonymisierte Daten, die nur in besonderen Fällen, evtl. mit hohem Aufwand und nicht unbedingt für alle Datensätze, wieder Einzelpersonen zugeordnet werden können. Aufgrund dieser Anforderungen sind sie dennoch nicht als «indirekt personenbezogen» einzuordnen. Absolut anonyme Daten fallen jedenfalls vollständig aus dem Datenschutz heraus. Für relativ anonyme Daten wäre jedoch zusätzlich zu prüfen, ob der konkrete Auftraggeber entsprechendes Zusatzwissen, wenn auch nur für einen Teil der Datensätze, besitzt, sodass sie für ihn als personenbezogen zu gelten hätten[22].

Zur Definition von Anonymität siehe weiters § 66 GentechnikG[23], wonach eine Probe selbst dann als anonym gilt, wenn sie ohne Namen nur mit einem Code versehen ist, der ausschließlich in der jeweiligen Einrichtung mit dem Namen verbunden werden kann. Datenschutzrechtlich handelt es sich daher um indirekt personenbezogene Daten, sodass die «Anonymität» explizit gesetzlich festgelegt werden musste (und mit der hier vertretenen Definition nichts zu tun hat).

### 4.2. Einsatz von Anonymisierung

Grundsätzlich ist kein Verbot, Anonymisierung selbst durchzuführen, in Österreich bekannt. Dies ist u.A. daran zu ersehen, dass z.B. das Vermummungsverbot bei Versammlungen (§ 9 Versammlungsgesetz), auch eine Art

[18] Richtlinie 95/46/EG des Europäischen Parlaments und des Rates vom 24. Oktober 1995 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr., ABl. L 281/31 vom 23. November 1995.

[19] Siehe auch Jahnel, Datenschutzrecht[2], RZ 3/84, 85.

[20] Vgl dazu oben die Datenanonymisierung durch Hashing.

[21] BGBl. Nr. 328/1968 i.d.F. BGBl. I Nr. 112/2007.

[22] Die genauen Konsequenzen einer derartigen Einteilung wären noch auszuarbeiten; so würde dies z.B. bedeuten, dass ein Auskunftsrecht besteht (da eine Identifizierung diesem Auftraggeber zumindest teilweise möglich ist), dieses jedoch öfters fehlschlagen wird (uU sind nicht aller Personen identifizierbar). Umgekehrt wäre eine Veröffentlichung jederzeit möglich, da die Daten für Dritte anonym wären.

[23] BGBl. Nr. 510/1994 i.d.F. BGBl. I Nr. 92/2015.

der Anonymisierung, explizit eingeführt wurde[24] und unter Vorbehalten/Einschränkungen steht. Ebenso legt § 40d Abs. 2 BWG[25] in Abweichung von der normalen Regelung fest, dass Kreditinstitute keine anonymen Konten führen dürfen. Vergleiche auch § 5 SignaturG, das Pseudonyme explizit vorsieht, bzw. die eIDAS-VO[26]. Letztere sieht ebenso Pseudonyme explizit vor (aber ErwG. 33: Mitgliedsstaaten dürfen eine Identifizierung verlangen – dies ist daher als begründete Ausnahme zu sehen) bzw. untersagt deren Verbot bei elektronischen Transaktionen (Art. 5 Abs. 2). Ein weiteres Beispiel ist das Anonymgeschäft an der Börse (§ 63 BörseG[27]).

Ganz im Gegenteil wird Anonymisierung teilweise sogar gesetzlich vorgeschrieben (z.B. § 99 TKG: Verkehrsdaten nach der Beendigung der Verbindung/Bezahlung oder in § 117b Abs. 1 Z 21 Ärztegesetz[28]: Anonyme Veröffentlichung von Fort-/Weiterbildungen auf der Homepage der Ärztekammer). Auch die Gewerbeordnung geht von der grundsätzlichen Möglichkeit aus, Geschäfte anonym durchzuführen, da nur in besonderen Fällen zusätzliche Maßnahmen erforderlich sind (z.B. § 365s Abs. 4 GewO[29]). Ähnlich dazu die Erläuterungen zur Konvention des Europarates zur Computerkriminalität in RZ 62[30], wonach die Anonymisierung von Kommunikationsdaten (Anonyme Remailer werden explizit erwähnt; ebenso Verschlüsselung) grundsätzlich als legitime Maßnahmen zur Sicherung der Privatsphäre angesehen werden und damit rechtmäßig sein sollen. Ausgenommen davon wäre gegebenenfalls die Veränderung von Daten («z.B. Paket-Header-Informationen», beträfe etwa die Anonymisierung von IP-Adressen) zur Verschleierung der Identität bei der Begehung einer Straftat, d.h. im Umkehrschluss wäre solches allgemein erlaubt.

Nicht alle direkt als Gesetz gelten diverse andere Rechtsgrundlagen, die Anonymität erlauben, z.B. Art. 12 der Allgemeinen Erklärung der Menschenrechte[31], der gleichlautende Art. 17 des Internationalen Pakts über bürgerliche und politische Rechte[32], Art. 8 EMRK/Art. 7 EU-GRC (Achtung des Privatlebens) oder Art. 10 EMRK/Art. 11 EU-GRC – Meinungs- und Kommunikationsfreiheit). Diese beziehen sich sehr allgemein auf Privatsphäre bzw. Datenschutz, sodass ein spezifisches Recht auf Anonymität im Internet bzw. die Verwendung von Anonymisierungstools daraus höchstens durch umfangreiche Auslegung ergibt. Weiters stehen diese unter Eingriffsvorbehalt und es bestehen ebenso gegenteilige Grundrechte (z.B. Art. 8 EMRK/Art. 7 EU-GRC – Achtung des Privatlebens im Hinblick auf die Verfolgung von Verstößen, Art. 5 StGG/ Art. 17 EU-GRC – Unversehrtheit des Eigentums in Bezug auf die Verfolgung von Online-Delikten). Ein gesetzliches «Verbot» für konkrete Fälle (d.h. nicht generell wie z.B. bei der Vorratsdatenspeicherung) kann daher u.U. verhältnismäßig, geeignet etc. sein. Zu den grundrechtlichen Aspekten sowie zu möglichen Einschränkungen der Rechte siehe den Report des UN-Menschenrechtsrates[33].

---

24 BGBl. I Nr. 127/2002.

25 BGBl. Nr. 532/1993 i.d.F. BGBl.I Nr. 108/2007.

26 Verordnung (EU) Nr. 910/2014 des Europäischen Parlaments und des Rates vom 23. Juli 2014 über elektronische Identifizierung und Vertrauensdienste für elektronische Transaktionen im Binnenmarkt und zur Aufhebung der Richtlinie 1999/93/EG, ABl. L 257/73 vom 28. August 2014.

27 BGBl. Nr. 555/1989 i.d.F. BGBl. Nr. 529/1993.

28 BGBl.I Nr. 169/1998 i.d.F. BGBl.I Nr. 56/2015.

29 Abs. 4: «Die Gewerbetreibenden haben der Gefahr der Geldwäsche oder Terrorismusfinanzierung aus Produkten oder Transaktionen, die die Anonymität begünstigen können, besondere Aufmerksamkeit zu widmen und allenfalls Maßnahmen zu ergreifen, um einem Missbrauch in dieser Hinsicht vorzubeugen.»: Anonyme Transaktionen sind daher selbst bei Gefahr erlaubt, sofern entsprechende Maßnahmen gegen Missbrauch eingesetzt werden.

30 Explanatory Report zur Convention on Cybercrime, http://conventions.coe.int/Treaty/en/Reports/Html/185.htm.

31 A/RES/217, UN-Doc. 217/A-(III).

32 BGBl Nr. 591/1978.

33 Kaye, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 22. Mai 2015, A/HRC/29/32, http://www.refworld.org/docid/5576dcfc4.html; siehe auch schon früher La Rue, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 16. Mai 2011, A/HRC/17/27, sowie La Rue, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 17. April 2013, A/HRC/23/40.

Umgekehrt besteht keine generelle Pflicht, anonyme Kunden zuzulassen. Websites oder beliebige Dienste im Internet können daher (wenn dies wirtschaftlich oder aus sonstigen Gründen sinnvoll erscheint) auf Identifikation beliebiger Verlässlichkeit aller Interessenten/Kunden bestehen. Analog dazu ist der generelle Ausschluss von Website-Besuchern zu sehen, die über Tor surfen und daher «anonym» (verschleierte IP-Adresse) sind. Siehe jedoch in Deutschland § 13 Abs. 6 TelemedienG[34].

Potentiell problematisch könnte § 295 StGB (Beweismittel-Unterdrückung) für Anonymisierung sein: Wird ein Datensatz vorsätzlich deshalb anonymisiert, damit er als Beweismittel nicht mehr geeignet ist, so ist dies strafbar. Dies wird dadurch eingeschränkt, dass sich ein – wenn auch nicht ausdrücklich erklärter – maßgebender Wille (wessen bleibt unklar bzw. ist sehr weit zu fassen) für die Verwendung in einem Verfahren entschieden hat und der Anonymisierende diesen Entschluss kennt oder doch mit ihm rechnet[35]. Weiters darf keine Alleinverfügungsberechtigung über die Daten bestehen. Meistens wird nicht beides auf Anonymisierung zutreffen, sodass das Delikt nicht verwirklicht wird[36]. Insb. für «Täter» ist OGH 11. Juni 2013, 14 Os 53/13w, relevant, wonach die Verhinderung der Beweismittelentstehung straflos ist. Wird Anonymisierung daher «vorbeugend» eingesetzt, um Spuren gar nicht entstehen zu lassen, ist dies nicht strafbar. Andere Verbote sind ebenso wenig ersichtlich. Gefährlich wird es, wenn fremde (Alleinverfügungsberechtigung) Daten nach Bekanntwerden (Verwendungswille) des Delikts anonymisiert oder gelöscht werden.

Ist die Verwendung von Anonymisierungsdienstes ein ausreichender Verdacht für eine Straftat? M.M. ist dies eindeutig zu verneinen, denn es existieren viele gute und legale Gründe hierfür. Besteht allerdings aus anderen Gründen ein Verdacht, so kann die Verwendung u.U. als verdachtsverstärkend angesehen werden. Denn allgemein ist die Verwendung entsprechender Software wenig verbreitet und zumindest derzeit ungewöhnliches Verhalten[37]. Daher kann es ein zusätzlicher Mosaikstein sein – selbst wenn später eine vollkommen korrekte und verständliche Alternative präsentiert wird. Anonymisierung kann daher zwar keinen Verdacht begründen, aber einen bestehenden erhöhen, sofern sie im konkreten Zusammenhang relevant ist. Allerdings stellt sich die Frage, wie dies zu erkennen ist: Ohne Abhören der Kommunikation oder Untersuchung des Computers ist dies nicht feststellbar. Und hierfür ist ohnehin bereits ein Verdacht erforderlich. Es verbleiben daher nur Aussagen und direkte Beobachtungen als mögliche legale Quellen.

### 4.3. Anbieten von Anonymisierungsdiensten

Für Mitwirkende an Anonymisierung, z.B. Betreiber eines Tor-Servers oder eines anonymen Remailers, gilt ähnliches wie für Nutzer: Sofern ihnen nicht bekannt ist, dass bereits Ermittlungen durchgeführt werden, kommt eine Strafbarkeit als Beitragstäter nicht in Betracht[38]. Auch zur Anstiftung für Straftaten sind sol-

[34] Telemediengesetz vom 26. Februar 2007, BGBl. I S. 179, i.d.F. BGBl. I S. 1324 vom 17. Juli 2015: § 13 Abs. 6 «Der Diensteanbieter hat die Nutzung von Telemedien und ihre Bezahlung anonym oder unter Pseudonym zu ermöglichen, soweit dies technisch möglich und zumutbar ist. Der Nutzer ist über diese Möglichkeit zu informieren.»

[35] Plöchl/Seidl in WK2 § 295 StGB (Stand: 1. September 2010, rdb.at) RZ 4; siehe auch Schwaighofer, Die Beweismittelunterdrückung nach § 295 StGB – Versuch einer erträglichen Auslegung, ÖJZ 1995, 376.

[36] Im anfangs angeführten Fall der Seitensprung-Webseite gab es ein Gerichtsverfahren, ob die Website betrügerischer Weise eine zu hohe Anzahl an weiblichen Mitgliedern angab – hier wäre eine Anonymisierung der Kundendatei durch das Unternehmen («zum Schutz der Kunden») daher strafbar (Aufgrund der Vertragsbeziehung zu Kunden besteht vermutlich keine Alleinverfügungsberechtigung; die AGBs wären jedoch genau zu prüfen). Rayner, Ashley Madison employee «told to create hundreds of fake profiles of alluring women», http://www.telegraph.co.uk/news/11817155/Ashley-Madison-employee-told-to-create-hundreds-of-fake-profiles-of-alluring-women.html.

[37] Aus diesem Grund ist der Einsatz von Verschlüsselung kein entsprechendes Merkmal (mehr), da heute praktisch jeder (bewusst oder unbewusst) Verschlüsselung verwendet und entsprechende Tools weite Verbreitung finden (z.B. Sicherung von Archiven wie ZIP-Dateien mit einem Passwort).

[38] Vergleiche dazu Reindl-Krauskopf, Computerstrafrecht im Überblick, 123 zur Kollision der Löschungsverpflichtung nach § 16 ECG mit § 295 StGB: Auch dort wird eine Strafbarkeit nur dann angenommen, wenn dem Host-Provider im Löschungszeitpunkt klar ist, dass die zu löschenden Daten in einem konkreten Verfahren als Beweismittel verwendet werden sollen (und nicht nur «könnten»!).

che Einrichtungen nicht bestimmt, denn der Einsatz in legaler Weise ist real existierend, sinnvoll und kein Feigenblatt. Erst eine entsprechende Bewerbung für illegale Tätigkeiten käme hierfür (nicht jedoch für § 295 StGB; siehe oben) in Frage[39].

Sowohl bloßer Betrieb wie Nutzung eines Anonymisierungsdienstes sind daher legal. Wird ein Anonymisierungsdienst bei einer Straftat eingesetzt, so ist auch dies (abgesehen von der Straftat selbst) nicht strafbar. Es könnte jedoch strafverschärfend sein, siehe z.B. § 33 Abs. 1 Z 6 StGB («heimtückisch»). Dies wird zwar angenommen wenn die Tat heimlich begangen wird, setzt allerdings besonderen Vertrauensbruch voraus[40]. Letzterer fehlt jedoch meist beim Einsatz von Anonymisierung. Allerdings könnte dies relevant sein, wenn jemand (z.B. basierend auf der IP-Adresse) wegen Vorfällen vom Betreiber eines Servers ausgesperrt wird, und nun mit solcher Hilfe zurückkehrt, um damit die Sperre zu umgehen und eine Straftat gegen den Serverbetreiber zu begehen. Allerdings fehlt auch hier m.M. nach ein besonderes Vertrauensverhältnis.

### 4.4. Anonymisierung fremder Daten zur Speicherung

Sollen Daten die bislang personenbezogen vorlagen anonymisiert werden, wird hiermit der Personenbezug endgültig entfernt. Es ist daher niemandem, insb. nicht dem Betroffenen, möglich, die Daten wieder jemandem/sich zuzuordnen. Dies bedeutet daher einen «Verlust» für alle Parteien. Verschiedene Personen können jedoch ein Interesse am Fortbestand besitzen (abgesehen von gesetzlichen Löschungsverboten, z.B. zu Dokumentationszwecken § 27 Abs. 3 DSG). Da allgemein Dritte keinerlei Rechte an den Daten besitzen, können sie vernachlässigt werden. Relevant sind jedoch solche Dritte, für die derzeit anderes gilt oder in Zukunft gelten könnte. Beispiele hierfür sind Behörden (z.B. interessiert das Finanzamt dass Rechnungen personenbezogen bleiben, um die Lieferung bzw. das Gegenstück des anderen Vertragspartners überprüfen zu können –Rechtsvorschriften) oder Personen, die in naher Zukunft einen Prozess gegen/mit dem Dateninhaber führen wollen und hierfür die Daten als Beweis benötigen (derzeit kein Rechtsanspruch). Im zweiten Falle wäre zu prüfen, ob eine Sorgfaltspflicht zur Datenerhaltung besteht. Diese kann sich aus konkretem Wissen ergeben, z.B. wenn man von der beabsichtigten Klagsführung weiß (siehe dazu oben – Beweismittelunterdrückung), oder aus Treuepflichten, z.B. gegenüber Vertragspartnern.

Abgesehen von Dritten ist der Betroffene selbst zu berücksichtigen. So besitzt er z.B. ein eigenes Interesse daran, dass die Dokumentation seiner Gesundheitsbehandlungen erhalten bleibt. Daraus kann jedoch kein generelles Individualrecht abgeleitet werden. Dies ist u.A. daraus zu schließen, dass kein Anspruch darauf besteht, mit bestimmten Daten in einer DB enthalten zu sein[41]. Umso weniger besteht daher ein Recht darauf, überhaupt darin vorzukommen (auch wenn viele Betreiber gerne zusätzliche Personen/Daten aufnehmen). Möchte der Betroffene daher die Daten in Bezug auf ihn selbst erhalten, so ist er auf eine gesetzliche oder vertragliche Verpflichtung hierfür angewiesen oder muss sich die Daten, z.B. über ein Auskunftsbegehren, verschaffen und selbst archivieren.

### 4.5. Datenlöschung durch Anonymisierung?

Einer Löschverpflichtung (§ 27 Abs. 1 DSG) könnte evtl. durch Anonymisierung entsprochen werden. Durch das Löschen soll der Auftraggeber permanent und irreversibel die Daten verlieren (bei bloß «vorübergehendem» Verlust, z.B. durch Anmerkungen, liegt nur eine «Sperre» vor). Dies erfordert nicht zwangsweise phy-

Dass die Prüfpflicht zeitlich knapp bemessen ist («unverzüglich»), verstärkt noch die erforderliche Eindeutigkeit.

39 Siehe dazu Bergauer, Das Betreiben eines Anonymisierungsdienstes im Internet als strafbarer Beitrag zur Verbreitung von Kinderpornografie? JusIT 5/2014, 161 (zu LG für Strafsachen Graz 30. Juni 2014, 7 Hv 39/14p).)

40 Ebner in WK2 § 33 StGB (Stand: 1. September 2014, rdb.at) RZ 20.

41 OGH 1. Oktober 2008, 6 Ob 195/08g: Sollen in einer Bonitätsdatenbank aufgrund eines Verlangens des Betroffenen bestimmte Daten gelöscht werden, so besteht kein Recht, dass die restlichen Daten enthalten bleiben; der Betreiber kann auch diese löschen.

sische Datenvernichtung, da z.B. bei Papierakten eine Schwärzung der entsprechenden Teile[42] ausreicht, bzw. eine vollständige Löschung der Eintragungen in Protokollbüchern[43], selbst wenn der eigentliche Akt weiterhin existiert[44] (damit liegen weiterhin personenbezogene Daten vor, da sich die Identität des Betroffenen aus dem Akt ergibt, selbst wenn dieser nur mehr schwer auffindbar ist und insb. nicht länger über den Namen des Betroffenen). Wird daher für *jedermann* der Personenbezug *vollständig* und *permanent* entfernt, so entspricht eine (im obigen Sinne absolute) Anonymisierung der Löschung[45]. Dies u.A. deshalb, da es sich dann nicht mehr um personenbezogene Daten handelt und sie daher aus dem Anwendungsbereich des DSG herausfallen. Besonders zu beachten ist jedoch, dass keine Wiederherstellbarkeit möglich sein darf (d.h. keine bloß relative Anonymität): So ist z.B. der erste und einzige anonymisierte Datensatz leicht einer bestimmten Person zuzuordnen. Ebenso ist auf Zusatzinformationen Dritter zu achten, sodass eine Reduktion von «direktem» auf «indirekten» Personenbezug keine Anonymisierung ist und daher ebenso wenig eine Löschung. Dementsprechend ist § 26 Abs. 7 DSG zu berücksichtigen, wonach ab dem Zeitpunkt der Kenntnis eines Auskunftsverlangens die Daten nicht mehr vernichtet werden dürfen. Dies gilt daher gleichermaßen für Anonymisierung, da hiermit das Auskunftsrecht in identischer Weise umgangen werden könnte («Keine Daten über Sie mehr aufzufinden»).

### 4.6. Anonymisierung fremder Daten zur Verarbeitung

Nach dem Datenschutzgesetz dürfen Daten zur Erstellung von Statistiken verwendet werden, wenn das Ergebnis anonym, d.h. nicht mehr auf Einzelpersonen zurückführbar, ist (§ 46 Abs 1 DSG). Entgegen Jahnel[46] ist m.M. nach nicht erforderlich, dass die Statistik «wissenschaftlich» sein muss[47], da sonst bei der Wendung «Für Zwecke wissenschaftlicher oder statistischer Untersuchungen …» die Statistik nicht erwähnt hätte werden müssen. Siehe auch ErwG. 29 sowie Art. 6 Abs. 1 lit. b der DS-RL, die ebenfalls nur gleichberechtigt sagt, dass Daten für «historische, statistische oder wissenschaftliche Zwecke» verwendet werden. Daher ist u.A. bei legalem Datenbesitz eine Anonymisierung erlaubt, sofern das Ergebnis eine anonyme Statistik darstellt. Im Hinblick auf Anonymisierungsdienste bedeutet dies, dass die Erstellung von Statistiken über ihre Nutzung legal ist, da die Daten legal besessen werden. Denn schließlich ist es Aufgabe des Dienstes, diese inhaltlich «umzuarbeiten». Darüber hinaus kommt nicht die Z 3 von § 46 Abs. 1 DSG in Frage, da z.B. IP-Adressen Dritter für den Anonymisierungsdienst nur indirekt personenbezogen sind (diese Ausnahme gilt daher z.B. nicht bei der inhaltlichen Anonymisierung von E-Mails, da dort – bis zum Ersetzen – noch die vollständige E-Mail Adresse des Absenders enthalten ist!).

[42] Eintragungen unleserlich machen reicht bei manuellen Dateien (Papier DSK 25. Juni 2004, K120.877/0017-DSK/2004.

[43] DSK 22. Oktober 2008, K120.857/0010-DSK/2008: Die Karteikarte wurde physisch vernichtet, im Protokollbuch jedoch lediglich der Eintrag geschwärzt.

[44] DSK 6. Februar 2008, K121.127/0003-DSK/2008. Siehe auch OGH 13. September 2012, 6 Ob 107/12x: Eine Archivierung ist keine Löschung: «Maßgeblich ist, dass der Auftraggeber auf die Daten wieder zugreifen und diese rekonstruieren kann.»

[45] Unklar und evtl anderer Meinung Jahnel, Datenschutzrecht[2], RZ 3/112, 159. Klar ist, dass eine bloße Erschwerung/Verhinderung der Auffindbarkeit keine Löschung darstellt – dies ist jedoch auch keine Anonymisierung.

[46] Jahnel, Datenschutzrecht[2] RZ 8/13, 439; identisch ErläutRV zu § 46 1613 d.B. (XX. GP).

[47] Auch die Verwendung «wissenschaftlicher Methoden» ist kaum zu rechtfertigen, da schon die bloße Definition äußerst schwierig wäre. Weiters ist es bei anonymen Ergebnissen irrelevant, ob sie richtig sind oder wenigstens nach einem richtigen Vorgehen erstellt wurden: Eine Gefahr für Einzelpersonen besteht nicht mehr.