# BIG DATA ANALYTICS UND ARTIKEL 11 DSGVO

Jakob Zanol / Felix Schmautzer

Jakob Zanol, Wissenschaftlicher Projektmitarbeiter/Managing Scientist, Universität Wien, Arbeitsgruppe Rechtsinformatik
Schottenbastei 10-16/2/5, 1010 Wien, AT
Jakob.Zanol@univie.ac.at

Felix Schmautzer, Wissenschaftlicher Projektmitarbeiter, Universität Wien, Arbeitsgruppe Rechtsinformatik
Schottenbastei 10-16/2/5, 1010 Wien, AT
Felix.Schmautzer@univie.ac.at

***Schlagworte:*** ***Datenschutzrecht, Big-Data-Analytics, k-anonymity, Art. 11 DSGVO***

***Abstract:*** *Trotz der steigenden Bedeutung personenbezogener Daten in einer datengetriebenen Wirtschaft gibt es auch zahlreiche Bestrebungen, deren Verarbeitung datenschutzfreundlich und datenminimierend zu gestalten – nicht zuletzt, um den Anforderungen der DSGVO nachzukommen. Dennoch führt die weite Auslegung des „Personenbezuges" i.S.d. Art. 4 Z. 1 DSGVO dazu, dass viele der gesetzten Maßnahmen (z.B. k-anonymity) nicht immer „echte" Anonymisierung von Daten (Beseitigung des Personenbezuges) erreichen. Dieser Aufsatz zeigt auf, warum diese Maßnahmen dennoch (mit Blick auf Art. 11 DSGVO) wesentliche Erleichterungen für Verantwortliche im Bereich Big-Data-Analytics bringen können.*

## 1. Einleitung

Obwohl der terminologische Ursprung von „Big Data" nun schon länger zurückliegt, so ist die Definition in der Praxis trotz oftmaliger Verwendung immer noch nicht scharf abgegrenzt. Bekannte Definitionen wie von Merv[1] und Gartner[2] reichen für juristische Zwecke zur Beschreibung analytischer Informationssysteme jedoch meist aus, in welchen Big Data sich von klassischen Datenoperationen insbesondere durch *Volume*, *Velocity* und *Variety* unterscheidet und zu seiner Verwaltung nichttriviale Hardware und besondere Datenbanksysteme benötigt. Die hier relevanten datenschutzrechtlichen Aspekte belaufen sich auf *Volume* und *Variety* und deren angepasste Verwaltungssysteme. So liefern Werbegiganten im Internet nun seit zwei Jahrzehnten Direktwerbung an jede Person, die in den letzten 20 Jahren das Web oder ein Smartphone bzw. ein Windows 10 Betriebssystem verwendet hat.[3] Doch auch andere, auf Sachdaten spezialisierte Verarbeitungen für Vorhersagemodelle wie demographische Entwicklung, Warenhandel und Personenverkehr[4] können durch alleinige Referenz zu anderen Datensätzen einen Personenbezug erhalten.[5]

Auch im Zeitalter der DSGVO[6] und ePrivacy-Regeln sollen jene und andere Branchen nichts zu befürchten haben, nunmehr sogar Privilegierung erfahren. Nicht selten handelt es sich dabei auch um Forschungspro-

---

1 Merv, A. It's going mainstream, and it's your next opportunity in: Teradata Magazine Online., https://bit.ly/3rJEt33 (abgerufen am 12. November 2021).

2 Beyer, M.: Gartner Says Solving "Big Data" Challenge Invoices More Than Just Managing Volumes of Data: https://bit.ly/3rJ6est (abgerufen am 12. November 2021).

3 Der digitale Direktwerbedienst Google Ads™ wurde bereits am 23. Okt. 2000 gegründet, vgl.: https://www.google.com/press/pressrel/pressrelease39.html (abgerufen am 12. November 2021); Windows 10 Datenschutzbestimmungen zur Werbe-ID in: https://bit.ly/3rK43F2 (abgerufen am 12. November 2021).

4 Dittmar, C. in: Gluchowski, P./Chamoni, P. (Hrsg.) Analytische Informationssysteme, Spinger, Berlin 2016, S. 59.

5 Weber, R./Oertly, D. Aushöhlung des Datenschutzes durch De-Anonymisierung bei Big Data Analytics?, Jusletter IT 21. Mai 2015, S. 4.

6 Verordnung (EU) 2016/679 ABl. L 2016/119, S. 1 (Datenschutz-Grundverordnung – DSGVO).

jekte, die vor allem im medizinischen Bereich große Signifikanz besitzen.[7] Nun, da sich der (Sand)sturm[8] rund um die DS-GVO gelegt hat und sich die Lehre sowie die Justiz und die übrigen Behörden ein Bild der Resultate machen konnten, ist es nicht nur an der Zeit Bilanz zu ziehen, sondern auch sich herauskristallisierte Besonderheiten zu thematisieren. Eine jener ist Art. 11 DS-GVO, welcher an sich keinerlei Schwesternbestimmung hat und lediglich mit den allgemeinen Bestimmungen in Art 5 lit e DSGVO in Verbindung steht.[9] Seine Einmaligkeit und seine Konsequenzen für die Praxis wurden jedoch in der Lehre bis dato vergleichsweise wenig ergründet. Diese sollen nun auf Basis eines Use Cases von Anonymisierung in Big Data dargestellt werden. Die Wirksamkeit von Anonymisierungsmethoden ist aufgrund potenzieller Verknüpfungen verschiedener Datensätze schwer vorherzusagen und noch schwerer gänzlich zu beseitigen. Daher hat die Aussage über den Personenbezug eines Datums uU eine durchaus geringe Halbwertszeit.

Somit soll Art 11 DSGVO und sein Einfluss in grundlegende Thematiken wie den Personenbezug selbst in den folgenden Ausführungen näher untersucht werden. Dabei wird spezifisch auf Anonymisierung durch Aggregierung bzw. auf k-Anonymität im Rahmen eines Beispiel-Datensatzes eingegangen.[10]

## 2. Personenbezug in Big (and long) Data

Wesentliches Kriterium für die Anwendung der DSGVO ist der Personenbezug. Der sachliche Anwendungsbereich der DSGVO ist gemäß Art 2 Abs. 1 DSGVO bereits dann eröffnet, wenn Verarbeitungsvorgänge „*personenbezogene Daten*“ umfassen.[11] Wann Daten „personenbezogen“ sind, wird in Art. 4 Z. 1 i.V.m. Erw. 26 DSGVO näher erläutert und die maßgeblichen Kriterien wurden in der Judikatur des EuGH bereits deutlich festgehalten. Dessen ungeachtet scheinen in der Literatur dennoch einzelne Punkte ungeklärt zu sein, was es rechtfertigt, die nach unserer Ansicht maßgeblichen Kriterien zur Prüfung, ob Daten „personenbezogen“ sind oder nicht, in Kürze darzustellen.

Einleitend wurde gesagt, dass das Kriterium des Personenbezugs iSd Art 4 Z. 1 DSGVO weit verstanden wird. Art 4 Z. 1 DSGVO enthält die Legaldefinition von „personenbezogenen“ Daten und lautet auszugsweise wie folgt: „*[...] alle Informationen, die sich auf eine* ***identifizierte oder identifizierbare*** *natürliche Person (im Folgenden ‚betroffene Person‘) beziehen [...]*“. In Erw 26 DSGVO wird zusätzlich festgehalten: „*Um festzustellen, ob eine natürliche Person* ***identifizierbar*** *ist, sollten alle Mittel berücksichtigt werden, die von dem Verantwortlichen oder einer anderen Person nach allgemeinem Ermessen wahrscheinlich genutzt werden, um die natürliche Person direkt oder indirekt zu identifizieren, wie beispielsweise das Aussondern.*“ An dieser Stelle ist zu erwähnen, dass Erw 26 DSGVO weitgehend gleichlautet wie Erw 26 der Datenschutzrichtlinie[12]. Zu der Rechtslage unter der DSRL hatte der EuGH in der Rechtssache Breyer entschieden, dass die zur Identifizierung erforderlichen Mittel (einschließlich Zusatzinformationen zur Identifizierung der betroffenen Person) nicht in den Händen einer einzigen Person befinden müssten, um eine „Identifizierbarkeit“ zu bejahen.[13] Im konkreten Fall entschied der EuGH, dass ein Webseitenbetreiber, der dynamische IP-Adressen der Besucher der Webseite

---

7 So nun sogar §2d Abs. 1 Z1 des Bundesgesetzes über allgemeine Angelegenheiten gemäß Art. 89 DSGVO und die Forschungsorganisation (Forschungsorganisationsgesetz – FOG), BGBl. Nr. 1981/341 idF BGBl. I Nr. 2021/205.

8 2021 sank das Publikationsvolumen zur DSGVO international, trotz steigender Zahlen in der Nutzung digitaler Medien im „Home Office“, um 25% (Elsevier Scopus, stand 31.12.2021) – unter jenes von 2018. Derselbe Trend zeichnet sich auch bei der DSB ab; siehe auch https://bit.ly/3GijE3l (abgerufen am 10.12.2021).

9 Wolff, H. in Wolff/Brink: BeckOK Datenschutzrecht, C.H. Beck, München 2021, Art. 11 Rz. 1.

10 Slijepcevic, D./Henzl, M./Klausner, L./Dam, T./Kieseberg, P./Zeppelzauer, M., k-Anonymity in Practice: How Generalisation and Suppression Affect Machine Learning Classifiers, Computers and Security 111, Elsevier, Wien 2021, S. 1.

11 Voll- oder teilautomatisierte Verarbeitungsvorgänge oder generell mit dem Ziel der Speicherung in ein Dateisystem; vgl. Art. 2 Abs. 1 DSGVO: „*Diese Verordnung gilt für die ganz oder teilweise automatisierte Verarbeitung personenbezogener Daten sowie für die nichtautomatisierte Verarbeitung personenbezogener Daten, die in einem Dateisystem gespeichert sind oder gespeichert werden sollen.*“

12 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, S. 31.

13 EuGH 19. Oktober 2016, C-582/14, Breyer, Rz. 43.

speichert, durchaus personenbezogene Daten speichert. Auch wenn dieser selbst sie keiner natürlichen Person zuordnen kann, reicht bereits der Umstand aus, dass der Webseitenbetreiber sie in rechtlich zulässiger Weise an eine Strafverfolgungsbehörde weitergeben kann, welche wiederum rechtliche Möglichkeiten hat, die dynamische IP-Adresse durch einen Internet Service Provider auf eine natürliche Person rückführen zu lassen. Damit zeigt sich, dass die „Identifizierbarkeit" einer natürlichen Person (Art. 4 Z. 1 2. Variante) sehr weit zu verstehen ist.[14] Hier ist auch anzumerken, dass der EuGH in seiner Entscheidung nicht auch eine tatsächliche Identifizierung der Person über die Strafverfolgungsbehörden voraussetzt. Bereits die Möglichkeit der Identifizierung über diesen plausiblen und rechtlich zulässigen Weg reicht dafür aus.[15]

Legt man dieses weite Verständnis des Kriteriums „personenbezogen" auf Big Data Vorgänge um, so zeigt sich, dass selbst dann, wenn eine Anonymisierung beabsichtigt ist, durchaus in einigen Fällen weiterhin personenbezogene Daten verarbeitet werden. Dies belegen auch äußerst rezente Entscheidungen der DSB und des österr. OGHs in welchen die Löschung eines Datenbankeintrags allein nicht ausreicht bzw. Wahrscheinlichkeitsaussagen bestehen bleiben.[16] Dies hat enorme Implikationen für Big-, aber auch „Long Data" Verarbeitungen, welche, durch Ausdehnung auf lange Zeiträume, Inferenzen von (und in) andere(n) Datenbanken zulassen und somit unter Sammlung von (vormals) reinen Sachdaten nun „*en masse*" Personenbezug erhalten.[17] Das Problem der Pseudo-Anonymität wurde weiters bereits 2014 von der Art.-29-Datenschutzgruppe als gefährlich eingestuft[18] und der Anknüpfungspunkt „K-Anonymität" in wissenschaftlichen Projekten behandelt.[19]

## 3. Anonymisierungstechniken und K-Anonymität (k-anonymity)

Um die Grenzen der gängigen Anonymisierungstechniken aufzuzeigen, bietet sich die Erklärung anhand eines Beispiels an. Hier wird auf jenes der K-Anonymität eingegangen. K-Anonymität und seine Derivate bedeuten, dass in einem Datensatz Quasi-Identifikatoren (QIDs) durch Aggregierung, Suppression bzw. Generalisierung eliminiert werden, sodass durch jene keine Rückschlüsse mehr auf natürliche Personen möglich sind.[20] Diese Methodik **eliminiert Personenbezug**, da Informationen in einem Datensatz daraufhin nicht mehr eindeutig zuordenbar sind. Wichtig ist diese Methode deswegen, da der Datensatz im Anschluss trotzdem zu einer weiteren (fast) **vollwertigen ML-Analyse** verwendet werden soll und kann. Um k-Anonymität zu gewährleisten, sollte es ausreichen, die QIDs wie erwähnt zu bearbeiten. Jene sind *in concreto* Attribute wie Standortinformationen (zB. PLZ), Alter, ethnische Zugehörigkeit oder Geschlecht, die (in Kombination mit externen Informationen) zur Re-Identifizierung von Personen verwendet werden könnten. Ein Datensatz selbst ist als k-anonym zu bezeichnen, wenn jede Information ihren Wert mit mindestens k-1 anderen Einträgen teilt.[21]

Zumeist wird in IT-Literatur zur K-Anonymität auf die anschließende Verwendbarkeit eines Datensatzes im ML eingegangen. In einer rezenten Untersuchung werden zum Zweck der Messung einer Re-Identifikation,

---

14 Eine einfachere Lösung fand die DSK in K213.180/0021-DSK/2013: Ambiguität aus mindestens 5 Personen.

15 Zitat: „*nach allgemeinem Ermessen wahrscheinlich*"; anders Waltraut Kotschy, Replik zu den Anmerkungen von Dietmar Jahnel zu den Auswirkungen der Vorabentscheidung des EuGH 19. 10. 2016, C-582/14 (Breyer) auf den Begriff der „personenbezogenen Daten" nach § 4 Z. 1 DSG 2000 in: Jusletter IT, 22.2.2017, S. 3.

16 DSB 16. 4. 2021, D124.2651 (unveröffentlicht); OGH 15.4.2021, 6 Ob 35/21x.

17 Siehe Risiken in Weber, R./Oertly, D., Jusletter IT 21. Mai 2015, S. 7.

18 Artikel-29-Datenschutzgruppe, Stellungnahme 05/2014 zu Anonymisierungstechnologien, WP 216, 10. April 2014, S. 13; zu den unterschiedlichen Zugängen der Artikel-29-Datenschutzgruppe in deren jeweiligen Leitlinien bei der Beurteilung des Personenbezuges („*objektiver*" oder „*subjektiver*" Ansatz) siehe Hötzendorfer, W., Datenschutz und Privacy by Design im Identitätsmanagement, OCG, Wien 2016, S. 10ff. sowie S. 44.

19 Schweighofer, E./Hötzendorfer, W./Varga, S., In: Cik, M., Fellendorf, M., Schweighofer, E., (Hrsg.), Rechtliche und ethische Aspekte der Echtzeitanalyse von Bewegungsströmen auf Basis von Daten aus Mobilfunk und sozialen Medien auf Großveranstaltungen (AGETOR), OCG 2016, Band 323, Wien 2016, S. 101.

20 Dazu bereits Tore Dalenius, Finding a Needle In a Haystack or Identifying Anonymous Census Records, Journal of Official Statistics, Vol.2, No.3, Stockholm 1986, S. 329.

21 Latanya Sweeney, k-anonymity: a model for protecting privacy. International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, 10 (5), World Scientific, Singapore 2002; S. 557–570.

vier Anonymisierungsalgorithmen analysiert, welche Generalisierung und Suppression verwenden, um einen Datensatz k-Anonym zu gestalten.[22] Die **Resultate variieren** erwartungsgemäß nach verwendetem Algorithmus, dem Datensatz und *k* selbst. Abstrakt lässt sich jedoch festhalten, dass **je höher *k***, **desto niedriger $F_I$**, was in einem höheren Grad an Anonymisierung resultiert. In Einzelfällen jedoch kann in einer Teilmenge, wie aber auch im vollständigen Datensatz bis hin zu *k=20*, die Re-Identifizierung durch die verwendeten Methoden aufgrund von Ähnlichkeiten in der verwendeten Datenstruktur sogar **erleichtert** werden.[23] Der stärkste Grad dieses Effektes konnte bei einer Generalisierung von kategoriellen (im Unterschied zu numerischen) Daten gemessen werden, mit welchen sich der De-Anonymisierungsalgorithmus leichter tat.[24]

Wie groß *k* nun konkret sein muss lässt sich mit Blick auf Entscheidungen der ehem. Österr. Datenschutzkommission sogar für die alte Rechtslage beantworten, was allerdings nach den obigen Ausführungen nicht mehr als haltbar erscheint.[25] Die DSB ließ unlängst in jenem Zusammenhang jedoch eine schlichte Suppression eines Eintrags genügen, wobei aber festzuhalten war, dass eine erfolgreiche Anonymisierung stark von Datensatz und Methodik abhängt und **nicht immer** gegeben ist.[26]

Werden jedoch mehrere Datensätze miteinander kombiniert oder besteht nur eine nach allgemeinem Ermessen wahrscheinlich eingesetzte Möglichkeit, Datensätze miteinander zu kombinieren, so ist dies bei der Beurteilung des Personenbezuges sogar noch stärker zu berücksichtigen. Je größer die Menge an Daten innerhalb der zu berücksichtigenden Datensätze, umso schwerer wird es, Daten darin zu anonymisieren bzw eine ReIdentifizierung zu verhindern. Grundsätzlich heißt das aber, dass in bestimmten Fällen **trotz Erreichung von k-Anonymität** in den einzelnen Datensätzen die Bestimmungen der **DSGVO voll zur Anwendung gelangen können**, weil die Daten weiterhin als personenbezogen gelten müssen (weil eine Rückführung auf eine identifizierbare natürliche Person nach allgemeinem Ermessen vorgenommen werden kann).

Eine Anonymisierung – oder immerhin eine Annäherung an eine solche bzw eine „weitgehende" Anonymisierung – kann in bestimmten Fällen geboten sein. Art 5 Abs. 1 lit b DSGVO sieht als wesentlichen Grundsatz für die Verarbeitung jenen der „Zweckbindung" vor. Personenbezogene Daten dürfen nur für „*festgelegte, eindeutige und legitime Zwecke erhoben werden*" und „*dürfen nicht in einer mit diesen Zwecken nicht zu vereinbarenden Weise weiterverarbeitet werden*".[27] Auf dem Grundsatz der Zweckbindung baut unter anderem jener der „Datenminimierung" auf, welcher vorsieht, dass personenbezogene Daten für diesen „*Zweck angemessen und erheblich sowie auf das für die Zwecke der Verarbeitung notwendige Maß beschränkt*" sind. Dieser Grundsatz der Datenminimierung wird noch durch die Verpflichtung des Verantwortlichen ergänzt, „*geeignete technische und organisatorische Maßnahmen* [zu treffen]*, die dafür ausgelegt sind, die Datenschutzgrundsätze wie etwa Datenminimierung wirksam umzusetzen und die notwendigen Garantien in die Verarbeitung aufzunehmen, um den Anforderungen dieser Verordnung zu genügen und die Rechte der betroffenen Personen zu schützen*".[28]

Damit besteht im Wesentlichen bereits die Verpflichtung des Verantwortlichen, nach Möglichkeit den Personenbezug der verarbeiteten Daten (auch durch die verwendete Technologie) zu reduzieren. Ein Verantwortlicher, der diesen Pflichten nachkommt, genießt im Regime der DSGVO durchaus wesentliche Vorteile, was maßgeblich durch folgende Bestimmung bedingt wird:

---

22 Slijepcevic, D./Henzl, M./Klausner, L./Dam, T./Kieseberg, P./Zeppelzauer, M., k-Anonymity in Practice: How Generalisation and Suppression Affect Machine Learning Classifiers, Computers and Security 111, Elsevier, Wien 2021, S. 1.

23 Slijepcevic et al., S.16.

24 Ibid., S. 26: insb. Abb. 9, siehe https://bit.ly/3EUzuiS (abgerufen am 12. November 2021).

25 DSK 22.5.2013, K213.180/0021-DSK/2013: „mehr als fünf Personen"; nun EG 26 DSGVO „*singling out*".

26 DSB 5.12.2018, DSB-D123.270/0009-DSB/2018.

27 Vgl. Art. 5 Abs. 1 lit b DSGVO.

28 Vgl. Art. 25 Abs. 1 DSGVO.

## 4. Artikel 11 DSGVO als Handreichung

Es wurde bereits mehrfach gezeigt, dass es in (weitgehend) automatisierter Form möglich ist, Datensätze in einer Weise zu aggregieren, mit der die Anforderungen, welche in der Rechtsprechung der nationalen Aufsichtsbehörde an eine echte, rechtliche Anonymisierung (d.h. Löschung des Personenbezuges) gestellt werden, grundsätzlich erreicht werden können.[29] Soeben wurden jedoch auch Gründe dargestellt, warum eine Anonymisierung eines Datensatzes (gerade in einem Big Data Kontext), in dem Datensätze stetig ergänzt und kombiniert werden, nicht dauerhaft garantiert werden kann.

Nach diesem Zwischenergebnis bietet sich jedoch aus rechtlicher Sicht ein Blick auf die Bestimmung des Art. 11 DSGVO an. Darin wird zunächst festgestellt, dass personenbezogene Daten, die ein Verantwortlicher verarbeitet, **um die betroffene Person zu identifizieren**, nicht bloß zum Zweck der Erfüllung allfälliger Pflichten (die eine solche Identifizierung voraussetzen) aufbewahrt oder sonst verarbeitet werden müssen.[30]

In Art. 11 Abs. 2 DSGVO wird darauf aufbauend festgestellt, dass in jenen Fällen, in denen eine Identifizierung nicht möglich ist,[31] die Artikel 15 bis 20 DSGVO keine Anwendung finden, sofern die betroffene Person nicht die dafür erforderlichen Informationen ergänzend zur Verfügung stellt. Zusätzlich sieht Art. 11 Abs. 2 DSGVO noch eine Informationspflicht bei fehlender Identifikationsmöglichkeit durch den Verantwortlichen vor (deren Umfang jedoch umstritten ist).[32]

Es zeigt sich somit, dass das maßgebliche Tatbestandsmerkmal des Art. 11 DSGVO jenes des „**Identifizierens**“ ist. Die Bedeutung desselben entscheidet die Reichweite der Bestimmung. Was also bedeutet „identifizieren“ i.S.d. Art. 11 DSGVO?

Zunächst wäre es naheliegend, den Begriff des „Identifizierens“ des Art 11 DSGVO als den Modus zur Herbeiführung des Personenbezuges i.S.d. Art. 4 Z. 1 DSGVO („**identifizierte oder identifizierbare** natürliche Person“) zu verstehen.[33] Damit stünde das Identifizieren in einem engen Zusammenhang mit dem „Personenbezug“ (iSv „personenbezogene Daten“). In der Literatur wird Hauptanwendungsbereich der Bestimmung in der Verarbeitung „pseudonymer“ oder „pseudonymisierter“ Daten beschrieben. „Pseudonymisierung“ i.S.d. Art. 4 Z. 5 DSGVO ist jedoch zu weit gefasst: während Art 11 Abs. 2 DSGVO Fälle betrifft, in denen eine „Identifizierung durch den Verantwortlichen“ nicht mehr möglich ist, ist dies bei der Pseudonymisierung nicht notwendig der Fall. Hat der Verantwortliche etwa eine separate Zuordnungstabelle[34], mit der den jeweiligen Pseudonymen eine natürliche Person zugeordnet werden kann, handelt es sich zwar um pseudonymisierte Daten, die Identifizierung durch den Verantwortlichen ist jedoch weiterhin möglich. Umgekehrt könnte Art. 11 DSGVO auch in Fällen zur Anwendung kommen, in denen Daten nicht eigens „pseudonymisiert“ werden: die oben thematisierte Leitentscheidung „Breyer“ ist dafür ein Beispiel.

Das **„Identifizieren“ i.S.d. Art. 11 Abs. 2 DSGVO** ist nicht als Identifizierung der betroffenen Person zu verstehen, sondern vielmehr als eine direkte **Zuordnung der verarbeiteten Daten** zu einem ganz **konkreten Antragsteller**: Dies wird in der Literatur z.T. auch aus dem Grund anerkannt, da in Art. 11 Abs. 2 DSGVO lediglich jene Betroffenenrechte (Art. 15–20 DSGVO) für unanwendbar erklärt werden, die einen Antrag der betroffenen Person voraussetzen (oder zumindest ermöglichen).[35] Nur dort macht es Sinn, deren Anwend-

---

29 Siehe oben, FN 9, S. 1.

30 Vgl. Art. 11 Abs. 1 DSGVO.

31 Hier besteht eine „Nachweispflicht“ des Verantwortlichen; vgl. etwa Klabunde, A., In: Ehmann, E./Selmayr, M., DS-GVO: Datenschutz-Grundverordnung: Kommentar[2], C.H. Beck/LexisNexis, München 2018, Art. 11 Rz. 16, 20.

32 Gola, P., In: Gola, P., DSGVO – Datenschutz-Grundverordnung: VO (EU) 2016/679: Kommentar[2], C.H. Beck, München 2018, Art. 11 Rz. 10 ist dies nur in Fällen möglich, in denen die betroffene Person sich zur Wahrnehmung von Rechten bei ihm meldet.

33 Vgl. hiezu Hötzendorfer, W., In: Knyrim, DatKomm – Praxiskommentar zum Datenschutzrecht – DSGVO, Manz, Wien 2020, Art. 11 Rz. 8 (rdb.at).

34 Beispiel von Jahnel, D., In: Jahnel, D., Kommentar zur Datenschutz-Grundverordnung (DSGVO), Jan Sramek, Wien 2021, Art. 11 Rz. 8.

35 Zu Art. 11 Abs. 2 DSGVO i.d.S. Feiler/Forgó, EU-DSGVO: EU-Datenschutz-Grundverordnung: Kurzkommentar (2017) Art. 11 Rz. 2 („*verarbeiteten Daten der antragstellenden Person zugeordnet werden [...]*“); so auch Klabunde, A., In: Ehmann, E./Selmayr, M., Art. 11 Rz. 19.

barkeit von der Möglichkeit der Identifizierung der betroffenen Person abhängig zu machen. Hat der Verantwortliche entsprechend dem Grundsatz der Datenminimierung seinen Datensatz soweit von Identifikatoren bereinigt, als dies im Rahmen des Verarbeitungszwecks möglich ist,[36] so soll ihm nicht vorgehalten werden, dass ein Antragsteller (ohne Zusatzinformationen) dem Datensatz nicht mehr ohne weiteres zugeordnet, d.h. „**als betroffene Person identifiziert**" werden kann. Dass aber eine betroffene Person ganz generell nicht mehr durch den Verantwortlichen identifiziert werden kann, ändert jedoch nichts an den grundsätzlichen Pflichten des Verantwortlichen bei der Verarbeitung von personenbezogenen Daten, einschließlich etwa der generellen **Löschpflicht** nach Art. 17 Abs. 1 DSGVO bei Wegfall des jeweiligen Erlaubnistatbestandes.[37]

Die begrifflichen Neuschöpfungen in der Literatur, wie etwa die „faktische Pseudonymisierung" (in Abgrenzung zur „rechtlichen Pseudonymisierung" i.S.v. Art. 4 Z. 5 DSGVO) oder der potentiell zu Missverständnissen führende Begriff der „relativen Anonymität",[38] können zwar beim Verständnis der Bestimmung begrenzt hilfreich sein, allerdings kann die damit einhergehende Vermengung der Konzepte des Personenbezuges (und dessen Fehlens) mit jenem der Identifizierung des Antragstellers i.S.d. Art. 11 DSGVO zur Verwirrung führen. Zumindest für Art. 11 Abs. 2 DSGVO sollten diese daher vermieden werden.

Daraus ergibt sich auch, dass die Möglichkeit der Identifizierung mittels Zusatzinformationen von Dritten (und auch von im Internet frei verfügbaren Informationen) für die Anwendbarkeit des Privilegs des Art. 11 Abs. 2 DSGVO keine Relevanz haben kann.[39] Art. 11 DSGVO steht auch nach der hier vertretenen Meinung in Zusammenhang mit der Obliegenheit des Antragstellers nach Art. 12 Abs. 2, 6 DSGVO seine Identität nachzuweisen.[40] Insoweit fügt sich die Obliegenheit des Antragstellers („betroffene Person") fehlende Informationen zu ergänzen gut in das System der DSGVO ein.

Das Ziel des Art. 11 DSGVO ist es somit, jenen Verantwortlichen, der bereits entsprechend den Grundsätzen der Zweckbindung sowie der (darauf aufbauenden) Datenminimierung und Speicherbegrenzung die verarbeiteten Informationen **auf ein Minimum reduziert** hat, davor zu schützen, ein Betroffenenrecht mangels Zuordenbarkeit zu dem jeweiligen Antragsteller verletzen zu müssen.[41] Angesichts dieses anerkannten Zwecks des Ausgleichs überrascht es, dass Art. 11 DSGVO dennoch sehr kritisch aufgenommen wurde und die Bestimmung zum Teil als in sich widersprüchlich[42] oder schlecht formuliert[43] oder auch als (angesichts des Datenminimierungsgrundsatzes) teilweise redundant[44] verstanden wird, deren Mehrwert fraglich[45] sei. Tatsächlich hat die Bestimmung aber durchaus eine ganz wesentliche Bedeutung, macht sie doch die in Art. 5 DSGVO geregelten **Grundsätze überhaupt erst praktikabel** (nach Jahnel hat Art. 11 DSGVO somit eine

---

36 Art. 11 Abs. 1 i.V.m. 5 Abs. 1 lit c („Datenminimierung").

37 Feiler/Forgó, EU-DSGVO: EU-Datenschutz-Grundverordnung: Kurzkommentar (2017) Art. 11 Rz. 2.

38 Gola, P., In: Gola, P., DSGVO, Art. 11 Rz. 5; dieser darf nicht verwechselt werden mit der „relativen Anonymität" wie sie Bergauer verwendet (demnach ist „relative Anonymität" dann erreicht, wenn eine Re-Identifizierung der betroffenen Person „nach allgemeinem Ermessen wahrscheinlich" [Erw. 26] nicht erfolgen wird).

39 So jedoch Weichert, T., In: Kühling, J./Buchner, B., Datenschutz-Grundverordnung/BDSG: Kommentar[2], C.H. Beck, München 2018, Art. 11 Rz. 13; dabei muss aber die Frage gestellt werden, wann ein solcher Fall vorliegen soll, wird doch ein Antragsteller im Internet leicht verfügbare Informationen ohnehin von sich aus mit dem Verantwortlichen teilen.

40 Siehe zu der entsprechenden Normenkonkurrenz zwischen Art. 11 Abs. 2 und Art. 12 Abs. 2 DSGVO insbesondere Veil, W., In: Gierschmann, S./Schlender, K./Stentzel, R./Veil, W., Kommentar Datenschutz-Grundverordnung, Bundesanzeiger Verlag, Köln 2018, Art. 11 Rz. 5 und 54 ff.; Art. 12 Abs. 2 DSGVO ist weiter formuliert als Art. 11 Abs. 2 DSGVO und umfasst auch Fälle des Art. 21 und 22 DSGVO; das Verhältnis dieser beiden Bestimmungen ebenfalls hervorhebend, Klabunde, A., In: Ehmann, E./Selmayr, M., DS-GVO, Art. 11 Rz. 21 sowie Paal, B./Hennemann, M., In: Paal, B./Pauly, D., Datenschutz-Grundverordnung, Bundesdatenschutzgesetz[3], C.H. Beck, München 2021, Art. 12 Rz. 46 ff.

41 Für viele etwa Gola, P., In: Gola, P., DSGVO, Art. 11 Rz. 4; Frenzel, In: Paal/Pauly, DSGVO Art. 11 Rz. 1.

42 Dabei den Unterschied der konkreten Identifizierung des Antragstellers und der generellen Identifizierbarkeit (s.o.) verkennend Georgieva, In: Kuner, C./Bygrave, L./Docksey, C./Drechsler, L., The EU General Data Protection Regulation (GDPR): a Commentary, Oxford University Press, Oxford 2020, Art. 11, S. 395.

43 Jahnel, D., DSGVO, Art. 11 Rz. 10 (Formulierung mehrfach misslungen); der zurecht a.a.O. in Rz. 16 („Potentiell betroffene Person" statt „betroffene Person"); Hötzendorfer, W., In: Knyrim, R., DatKomm (der „missglückte Wortlaut von Art. 11 Abs. 2 [...]").

44 Vgl. Frenzel, E., In: Paal, B./Pauly, D., DSGVO, Art. 11 Rz. 12.

45 Ibid. („keine Innovation"); Jahnel, D., DSGVO, Art. 11 Rz. 13 und 14.

„**Brückenfunktion**“[46]). Art. 11 DSGVO adressiert auch den Umstand, dass eine **Anonymisierung nicht immer erreicht** werden kann, ohne den Zweck der Verarbeitung zu negieren. Ohne Art. 11 DSGVO stünden Verantwortliche tatsächlich vor der Schwierigkeit, dass sich eine Identifizierbarkeit der betroffenen Person im Rahmen der beabsichtigten Tätigkeit zwar verringern, jedoch nicht zur Gänze ausschließen lässt. Das weite Verständnis des Personenbezuges (oben) hätte diesfalls zur Konsequenz, dass selbst eine konsequente Umsetzung des Grundsatzes der Datenminimierung und des Einsatzes von technischen Anonymisierungstechniken einen Personenbezug nicht (immer) ausschließt.

Dass sich Verantwortliche tatsächlich in einer solchen Zwickmühle befinden können, zeigt eine jüngere Entscheidung der Datenschutzbehörde sehr deutlich. Dieser Entscheidung zugrunde lag ein Gutachten zu „Fragen der Doppelansässigkeit in Österreich und der Schweiz“.[47] Dabei wurden der Beschwerdegegnerin Steuerinformationen zur Verfügung gestellt. Da sich darin jedoch im konkreten Fall „*bspw. Aussagen über den Wohnort, die familiäre und wirtschaftliche Situation, soziale Kontakte der Beschwerdeführerin sowie Interessen, Hobbies, die Zugehörigkeit zu Clubs oder einer Stammrunde zum einen direkt, zum anderen indirekt über den Bezug zum Ehemann der Beschwerdeführerin*“ ableiten ließen, waren diese „*grundsätzlich geeignet, einen Personenbezug zur Beschwerdeführerin [...] herzustellen*“. Im Ergebnis war sohin die Beschwerdegegnerin, die auf Basis der – ihrer Ansicht nach – anonymen Daten, zur Auskunft verpflichtet, was die Feststellung der Verletzung des Betroffenenrechts zur Folge hatte. Die DSB sah durchaus einen potenziellen Anwendungsfall des Art. 11 DSGVO, auf welchen sich die Beschwerdegegnerin jedoch nicht berufen hatte!

Diese Entscheidung verdeutlicht somit die oben hervorgehobene Dilemmasituation mancher Verantwortlicher. Hätte der Verantwortliche im gegebenen Fall weitere Informationen eingeholt, hätte er die Auskunft erteilen können. Da er jedoch lediglich die (vermeintlich) anonymen Daten verarbeitete, konnte er der betroffenen Person keine Auskunft geben, was grundsätzlich – da Art. 15 DSGVO keine entsprechende Ausnahme vorsieht – eine Verletzung dieses Rechts zur Folge hätte. Hier erfüllt Art. 11 DSGVO somit eine ganz wesentliche Funktion und erlaubt Verantwortlichen sohin die Datenminimierung in einem möglichst hohen Ausmaß, selbst wenn durch diese eine Anonymisierung allenfalls noch nicht erreicht wird/werden kann (und schützt vor der oben dargestellten Bredouille).

## 5. Fazit

Es wurde gezeigt, dass insbesondere im Kontext von Big Data Verarbeitungen eine Anonymisierung oft nicht (dauerhaft) erreicht werden kann und daher selbst bei der Anwendung von Anonymisierungstechniken ein Personenbezug der Daten im Lebenszyklus der Daten (wieder) bestehen kann. Dies ist dem weiten Verständnis von „personenbezogenen“ Daten geschuldet, welches mit der Entscheidung des EuGH in der Rechtssache Breyer nun anhand klarer Kriterien eindeutig definiert wurde. Damit kommt es aber zu der seltsamen Situation, dass die Anwendung von Anonymisierungstechniken, die zwar keine rechtliche Anonymisierung erreichen, aber immerhin die Verarbeitung der Daten auf das notwendige Minimum reduzieren (und damit die Grundsätze der Zweckbindung, Datenminimierung und Speicherbegrenzung umsetzen), den Verantwortlichen an der Erfüllung der Betroffenenrechte verhindern. Nach der Beseitigung des Klarnamens und verschiedener weiterer Kennungen aus einem Datensatz, kann der Verantwortliche die verbleibenden Daten mitunter nicht mehr einem konkreten Antragsteller zuordnen.

Die Betroffenenrechte sehen jedoch für diesen Fall keine Ausnahme vor – sie gelten grds. uneingeschränkt bei der Verarbeitung von personenbezogenen Daten in dem oben dargestellten weiten Verständnis. Diese Dilemmasituation wird auch anhand der hier dargestellten Entscheidungen der österreichischen DSB verdeutlicht, in denen „vermeintlich anonymisierte Daten“ Verantwortliche gerade in diese Dilemmasituation

[46] Siehe Jahnel, D., DSGVO, Art. 11 Rz. 2 („Brückenfunktion“).
[47] DSB 14.01.2019, DSB-D123.224/0004-DSB/2018.

manövrierten. Doch genau hier schafft Art. 11 DSGVO einen pragmatischen Ausgleich und stellt in eleganter Weise im Wesentlichen fest, dass die Zuordnung eines Datensatzes zum jeweiligen Antragsteller anhand der vorhandenen Informationen möglich sein muss, damit ein Betroffenenrecht wahrgenommen werden kann. Automatisierte Anonymisierungstechniken erlauben eine Reduzierung der Daten auf das für den jeweiligen Zweck notwendige Ausmaß, auch wenn dies bedeutet, dass notfalls die betroffene Person ergänzende Informationen zur Verfügung stellen muss, um dem Verantwortlichen die Wahrnehmung der Betroffenenrechte zu ermöglichen.

Art. 11 kann Verantwortlichen zwar eine Erleichterung bieten, die Bestimmung setzt jedoch eine Beachtung der Grundsätze der Verarbeitung nach Art. 5 DSGVO voraus. So darf keinesfalls eine Identifizierung des Antragstellers absichtlich verunmöglicht werden (Grundsatz von *Treu und Glauben*), oder gegen die Verpflichtung zur „Privatsphäre durch Technikgestaltung“ verstoßen werden.

## 6. Danksagung

This research was funded by the Austrian Research Promotion Agency (FFG) COIN project 866880 “Big Data Analytics”. The financial support by the Austrian Research Promotion Agency and the Federal Ministry for Digital and Economic Affairs is gratefully acknowledged.

Das Projekt wird von der Arbeitsgruppe Rechtsinformatik, Juridicum, Universität Wien, unter der Leitung von Prof. Dr. Dr. Erich Schweighofer durchgeführt. Die Autoren danken für die wesentliche Unterstützung und wichtige Hinweise.