

Kerstin Noëlle Vokinger

Gesundheitsdaten im digitalen Zeitalter

Gesundheitsdaten gelten rechtlich als besonders schützenswerte Daten. Der vorliegende Beitrag zeigt auf, dass als Folge des technologischen und medizinischen Fortschritts auch das rechtliche Verständnis, was «Gesundheitsdaten» sind und welche Daten tatsächlich als anonymisiert zu qualifizieren sind, einem Wandel unterliegt.

Beitragsart: Beiträge

Rechtsgebiete: Gesundheitsrecht; Datenschutz

Zitiervorschlag: Kerstin Noëlle Vokinger, Gesundheitsdaten im digitalen Zeitalter, in: Jusletter 27. Januar 2020

Inhaltsübersicht

1. Definition «Gesundheitsdaten» vor dem Hintergrund der technologischen Entwicklungen
 - 1.1. Von manuell erfassten Einzeldaten auf Papier mit wenigen Akteuren...
 - 1.2. ... zu digitalen Datenbanken mit zahlreichen Akteuren
 - 1.3. Zwischenfazit
2. «Anonymitätsgrad» von (Gesundheits-)Daten und dessen Problematik
 - 2.1. Definitionen
 - 2.2. Gibt es noch anonymisierte oder anonyme Daten?
3. Fazit

1. Definition «Gesundheitsdaten» vor dem Hintergrund der technologischen Entwicklungen

1.1. Von manuell erfassten Einzeldaten auf Papier mit wenigen Akteuren...

[1] Als Personendaten gelten alle Angaben, die sich auf eine bestimmte oder bestimmbare Person beziehen.¹ Der Gesetzgeber klassifiziert gewisse Personendaten als *«besonders schützenswert»*. Hierzu zählen, neben beispielsweise Daten über die religiösen oder politischen Ansichten, auch *Daten über die Gesundheit*.² Im Gegensatz zu den übrigen Daten, gelten für besonders schützenswerte Daten strengere Vorschriften. Beispielsweise muss die Einwilligung für die Bearbeitung von Gesundheitsdaten ausdrücklich erfolgen,³ private Personen müssen Datensammlungen beim Beauftragten zur Registrierung anmelden⁴ oder die Daten dürfen Dritten ohne Rechtfertigungsgrund nicht bekanntgegeben werden.⁵

[2] Auf eine Legaldefinition des Begriffs «Gesundheitsdaten» verzichtet der Gesetzgeber jedoch. Konkretisiert wird der Begriff in der Botschaft und in der Lehre. Demnach sind darunter alle Informationen zu verstehen, die Aufschlüsse über medizinische Befunde geben, die sich für die Betroffenen negativ auswirken können. Dabei ist es nicht notwendig, dass es sich um eine konkrete Diagnose handelt. Vielmehr reichen bereits Ergebnisse aus medizinischen Untersuchungen (etwa Anamnese, körperliche Untersuchungen oder labor- und gerätemedizinische Untersuchungen) sowie die in Patientenrechnungen enthaltenen Daten, da Letztere Rückschlüsse über den Gesundheitszustand des Patienten erlauben.⁶

[3] Das gleiche Ergebnis ergibt sich auch aus dem *Humanforschungsgesetz*, wonach gesundheitsbezogene Personendaten alle Informationen über eine bestimmte oder bestimmbare Person sind, die sich auf deren Gesundheit oder Krankheit beziehen, einschliesslich ihrer genetischen Daten.⁷

¹ Art. 3 lit. a DSG (Bundesgesetz über den Datenschutz vom 1. Juni 1992, SR 235.1).

² Art. 3 lit. c. DSG.

³ Art. 4 Abs. 5 DSG.

⁴ Art. 11a Abs. 2 und Abs. 3 DSG.

⁵ Art. 13 DSG.

⁶ Vgl. zum Ganzen Botschaft vom 23. März 1988 zum Bundesgesetz über den Datenschutz, BBl 1988 II 413 ff., 446; REGINA AEBI-MÜLLER/WALTER FELLMANN/THOMAS GÄCHTER/BERNHARD RÜTSCHKE/BRIGITTE TAG, *Arztrecht*, Bern 2016, S. 437; vgl. auch GABOR P. BLECHTA, Art. 3 DSG, Rz. 33, in: Urs Maurer-Lambrou/Gabor P. Blechta (Hrsg.), *Basler Kommentar Datenschutzgesetz / Öffentlichkeitsgesetz*, 3. Auflage, Basel 2014 (zit. BSK DSG-AUTORIN).

⁷ Art. 3 lit. f HFG (Humanforschungsgesetz vom 30. September 2011, SR 810.30).

[4] Deutlich wird, dass «Gesundheitsdaten» als wandelbarer Begriff zu verstehen ist, wobei die Umschreibungsmerkmale einen Ermessensspielraum gewähren, ob im Einzelfall Gesundheitsdaten vorliegen oder nicht. Dementsprechend kann der Kreis der Gesundheitsdaten nicht abschliessend aufgezählt werden.

[5] Traditionellerweise ergab sich der Grossteil der Gesundheitsdaten aus der Arzt-Patienten-Beziehung.⁸ Die entsprechenden Daten dienten primär als Grundlage für die Diagnose und Therapie des Patienten. Der Arzt erfasste die medizinischen Informationen manuell bzw. analog und bewahrte sie in Form von physischen Dossiers bzw. Krankengeschichten in der Arztpraxis oder im Spital auf. Auch weitere Akteure hatten Zugang zu (gewissen) Gesundheitsdaten, zu denken ist insbesondere an Wissenschaftler bzw. wissenschaftliche Institutionen, (Kranken-)Versicherungen oder den Arbeitgeber. Im Regelfall griffen aber auch diese Akteure auf die Daten zurück, die sich aus der Arzt-Patienten-Beziehung ergeben haben.

1.2. ... zu digitalen Datenbanken mit zahlreichen Akteuren

[6] Der allgemeine Fortschritt in der Informationstechnologie führte in den letzten Jahren dazu, dass in allen Lebensbereichen mehr Daten bearbeitet⁹ werden. Dies gilt auch für den Gesundheitssektor. So werden in den Spitälern und vermehrt auch in den ambulanten Praxen Patientendaten digital erfasst und in Datenbanken gespeichert.

[7] Neben den klassischen Akteuren sind in den letzten Jahren zunehmend auch *neue Akteure* in den Gesundheitsmarkt eingedrungen. Beispielsweise sammeln und analysieren grosse Unternehmen wie Google oder Amazon länderübergreifend Gesundheitsdaten, wobei diese Unternehmen um ein Vielfaches mehr Gesundheitsdaten sammeln und analysieren als Spitäler.¹⁰ Das Sammeln von Gesundheitsdaten erfolgt nicht mehr aus der Arzt-Patienten-Interaktion, sondern zunehmend über *digitale Kanäle*, wie etwa Apps, *direkt vom Patienten bzw. Konsumenten*. Auf diese Art und Weise sammeln nicht nur private Unternehmen, sondern beispielsweise auch (soziale) Krankenversicherungen direkt Gesundheitsdaten von ihren Versicherten, z.B. die Anzahl der zurückgelegten Schritte pro Tag.¹¹ Dabei ist es nicht immer eindeutig, ob es sich bei den Daten um Gesundheitsdaten oder um Lifestyle-Daten handelt. Letztere gelten rechtlich nicht als besonders schützenswert. Während es sich bei der Anzahl der zurückgelegten Schritte pro Tag bei einem gesunden Menschen zumindest zum Zeitpunkt der Datenerhebung kaum um einen medizinischen Befund im Sinne des Gesetzgebers handelt, kann die gleiche Information beispielsweise

⁸ Hierzu gehören im weiteren Sinn auch die rechtlichen Beziehungen, die zwischen Patienten und den Spitälern (oder anderen Institutionen, die Ärzte und anderes Gesundheitspersonal beschäftigen) entstehen (vgl. THOMAS GÄCHTER/BERNHARD RÜTSCHKE, Gesundheitsrecht, 4. Auflage 2018, S. 71).

⁹ «Bearbeiten» ist im Sinne von Art. 3 lit. e DSGVO zu verstehen, wonach unter diesem Begriff jeder Umgang mit Personendaten erfasst ist, insbesondere das Beschaffen, Aufbewahren, Verwenden, Umarbeiten, Bekanntgeben, Archivieren oder Vernichten.

¹⁰ Vgl. etwa <https://www.wsj.com/articles/google-s-secret-project-nightingale-gathers-personal-health-data-on-millions-of-americans-11573496790>; <https://www.theguardian.com/society/2019/dec/08/nhs-gives-amazon-free-use-of-health-data-under-alexa-advice-deal>.

¹¹ Vgl. etwa das Urteil des Bundesverwaltungsgerichts in Bezug auf das von der Helsana Zusatzversicherungen AG betriebene App-gestützte Bonusprogramm «Helsana+» (Urteil A-3548/2018, Medienmitteilung abrufbar unter <https://www.bvger.ch/bvger/de/home/medien/medienmitteilungen-2019/datenbearbeitung-bei-helsana---teilweise-rechtswidrig.html>).

bei Diabetikern oder herzinsuffizienten Patienten wichtige Informationen über ihren Gesundheitszustand bzw. Krankheitsverlauf geben und damit einen medizinischen Befund darstellen.

[8] Nicht nur die Abgrenzung zu Lifestyle-Daten muss im Einzelfall beurteilt werden. So können beispielsweise Informationen über Reisen je nach Fragestellung als Gesundheitsdaten oder Mobilitätsdaten qualifiziert werden. Handelt es sich um einen Patienten mit Infektionssymptomen (Erbrechen, Durchfall, Fieberschübe, Kopfschmerzen), so können Informationen über sein Reiseverhalten (etwa in afrikanische oder asiatische Länder) relevant sein für die Diagnosestellung. Bei gesunden Personen kann es sich demgegenüber bei diesen Informationen um «reine Mobilitätsdaten» handeln.

[9] An diesen Beispielen wird ersichtlich, dass der jeweilige *Kontext* der Datenbearbeitung relevant sein kann für die Beurteilung, ob Gesundheitsdaten vorliegen oder nicht.

[10] Als es einem Forscherteam im Rahmen des «Human Genome Project» im Jahre 2003 gelang, die menschliche DNA zum ersten Mal zu sequenzieren, d.h. die Reihenfolge aller Basenpaare zu determinieren, war die Geburtsstunde der Bioinformatik endgültig eingeläutet. Dieser technologische Fortschritt hatte nicht nur Konsequenzen für die Medizin, sondern damit einhergehend bildete dies eine weitere Dimension, wie *individuell-spezifische* Datendetails über eine Person gesammelt und analysiert werden können.¹² Darüber hinaus erlauben diese technologischen Fortschritte, dass Gesundheitsdaten nicht mehr nur auf bereits manifeste Erkrankungen reduziert werden, sondern durch die Erkennung von genetischen *Prädispositionen für bestimmte Erkrankungen* (z.B. Chorea Huntington oder die auf BRCA-1 oder BRCA-2 zurückführenden Mammakarzinome) bereits zu einem Zeitpunkt vor der klinischen Manifestation Gesundheitsdaten gesammelt werden.

1.3. Zwischenfazit

[11] Die technischen Entwicklungen und medizinischen Fortschritte in den letzten Jahren führen dazu, dass sich die Art und Anzahl von Daten, die als «Gesundheitsdaten» im rechtlichen Sinn zu qualifizieren sind, verändert haben. Diese Entwicklungen können wie folgt zusammengefasst werden:

- Gesundheitsdaten, die sich früher traditionell aus der Arzt-Patienten-Beziehung ergaben, werden heute zunehmend direkt vom Patienten bzw. vom Konsumenten direkt erfasst (etwa über Apps);
- neben den klassischen Akteuren im Gesundheitswesen bearbeiten vermehrt auch dem Gesundheitswesen bis vor kurzem fachfremde Akteure (z.B. Google, Amazon, Apple etc.) Gesundheitsdaten;
- zusätzlich zu den klinisch manifesten Daten werden vermehrt auch (genetische) Gesundheitsdaten, die über gewisse Prädispositionen Aufschluss geben, gesammelt.
- Die Abgrenzung zwischen Gesundheitsdaten und Lifestyle-Daten (oder anderen Datenkategorien) ist nicht immer klar. Zunehmend wichtig wird der jeweilige *Kontext* im Einzelfall. Die gleichen Daten (z.B. Anzahl zurückgelegter Schritte pro Tag) können bei einer (gesun-

¹² Vgl. statt vieler <https://www.genome.gov/human-genome-project/What>; International Human Genome Sequencing Consortium, Initial sequencing and analysis of the human genome, *Nature* 2001(409); 860–021.

den) Person als Lifestyle-Daten und bei einem Patienten als Gesundheitsdaten qualifiziert werden.

2. «Anonymitätsgrad» von (Gesundheits-)Daten und dessen Problematik

2.1. Definitionen

[12] Der Gesetzgeber unterscheidet zwischen verschiedenen Kategorien von Daten in Bezug auf ihren «Anonymitätsgrad». Ausgangspunkt sind die *Personendaten*. Hierzu gehören alle Angaben, die sich auf eine *bestimmte oder bestimmbare Person* beziehen.

[13] Eine Person gilt als bestimmt, wenn sich aus der Information selbst ergibt, dass es sich um diese konkrete Person handelt. Als Beispiel ist der Personalausweis aufzuführen. Demgegenüber gilt eine Person als bestimmbar, wenn die Möglichkeit besteht, ihre Identität festzustellen. Hierzu reicht es, wenn sich die Person aus dem Kontext der Informationen ermitteln lässt. Dies kann auf unterschiedliche Art erfolgen, beispielsweise anhand der AHV-Nummer, eines Aktenzeichens, einer Kundennummer oder eines Schlüssels. Dabei darf der für die Identifizierung erforderliche Aufwand nicht übermässig sein.¹³ Der Aufwand ist dann nicht mehr vertretbar, «wenn nach den allgemeinen Lebenserfahrungen nicht damit gerechnet werden muss, dass ein Interessent diesen auf sich nehmen wird.»¹⁴ Zu berücksichtigen sind in diesem Zusammenhang der *Stand der Technik* sowie die *technischen Entwicklungsmöglichkeiten*.¹⁵

[14] Werden personenbezogene Merkmale von Personendaten entfernt, sodass es danach niemandem mehr möglich ist, aus den Daten einen Personenbezug herzustellen, liegen *anonymisierte Daten* vor.¹⁶

[15] Demgegenüber werden bei einer Pseudonymisierung die personenbezogenen Merkmale nicht entfernt, sondern durch Pseudonyme ersetzt, d.h. einen Schlüssel, der nach einer bestimmten Regel den ursprünglichen personenbezogenen Merkmalen zugeordnet werden kann. Folglich stellen *pseudonymisierte Daten* nur für jene Personen weiterhin Personendaten dar, die den Schlüssel kennen. Für andere handelt es sich nicht mehr um Personendaten.¹⁷

[16] Anonyme Daten sind keine Personendaten mehr. Sie resultieren entweder aus anonymisierten Daten, die sich zuvor auf eine bestimmte oder bestimmbare Person bezogen haben, oder die Daten können bereits anonym erhoben werden.¹⁸

[17] Gerade aus rechtlicher Perspektive ist diese Unterscheidung wichtig, da anonymisierte Daten und anonyme Daten nicht unter die (Datenschutz-)Gesetzgebung fallen und entsprechend nicht denselben Schutz geniessen wie Personendaten.

¹³ Vgl. zum Ganzen BSK DSG-BLECHTA (Fn. 6), Art. 3 DSG, Rz. 9 f.; DAVID ROSENTHAL/YVONNE JÖHRI, Handkommentar zum Datenschutzgesetz, Zürich/Basel/Genf 2008, Art. 3 Rz. 36.

¹⁴ Botschaft vom 23. März 1988 zum Bundesgesetz über den Datenschutz, BBl 1988 II 413 ff., 445.

¹⁵ BSK DSG-BLECHTA (Fn. 6), Art. 3 DSG, Rz. 11.

¹⁶ ROSENTHAL/JÖHRI (Fn. 13), Art. 3 Rz. 35.

¹⁷ Vgl. zum Ganzen ROSENTHAL/JÖHRI (Fn. 13), Art. 3 Rz. 36.

¹⁸ BSK DSG-BLECHTA (Fn. 6), Art. 3 DSG, Rz. 13; ROSENTHAL/JÖHRI (Fn. 13), Art. 3 Rz. 3.

2.2. Gibt es noch anonymisierte oder anonyme Daten?

[18] Bereits seit längerer Zeit wird in der Informationstechnologie in Frage gestellt, ob es überhaupt noch anonyme Daten gibt.¹⁹ LATANYA SWEENEY, Professorin an der Universität Harvard, hat bereits während ihres Studiums im Jahre 1997 Studienergebnisse präsentiert, wonach es aufgrund drei relativ einfacher demographischer Merkmale – nämlich der Postleitzahl des Wohnortes, dem Geschlecht und dem Geburtstag – möglich ist, bei 87% der Bevölkerung in den USA die individuelle Person zu identifizieren.²⁰ Die Kombination von verschiedenen Daten und Datenbanken kann als «Linkage» umschrieben werden.

[19] Wir haben mit *Schweizer* Daten ebenfalls untersucht, ob eine Re-Identifikation von anonymisierten Daten möglich ist. Ausgangspunkt waren Gerichtsurteile, die häufig in anonymisierter Form öffentlich zugänglich gemacht werden. Materiell interessierten uns die Fragestellungen, welche pharmazeutischen Unternehmen zwischen 2000 und 2018 in einem Verfahren gegen (Preis-)Verfügungen des Bundesamtes für Gesundheit (BAG) in einem bundesgerichtlichen Verfahren involviert und welche Arzneimittel davon betroffen waren. Methodisch gingen wir so vor, dass wir verschiedene, öffentlich zugängliche Daten(banken) im Sinne des «Linkage» miteinander verbunden haben (Abbildung 1). Pro Bundesgerichtsentscheid wurde für die Re-Identifikation ein maximaler Zeitaufwand von einer Stunde aufgewendet. Basierend auf dieser Methodik erzielten wir eine Re-Identifikation in 84% der Fälle.²¹

¹⁹ Vgl. für die Schweiz ROLF H. WEBER/DOMINIC OERTLY, Aushöhlung des Datenschutzes durch De-Anonymisierung bei Big Data Analytics, Jusletter IT, 21. Mai 2015.

²⁰ LATANYA SWEENEY, Simple Demographics Often Identify People Uniquely. Data Privacy WorkingPaper, Pittsburgh 2000 (abrufbar unter <https://dataprivacylab.org/projects/identifiability/index.html>).

²¹ Vgl. zum Ganzen KERSTIN NOËLLE VOKINGER/URS JAKOB MÜHLEMATTER, Re-Identifikation von Gerichtsurteilen durch «Linkage» von Daten(banken). Eine empirische Analyse anhand von Bundesgerichtsbeschwerden gegen (Preisfestsetzungs-)Verfügungen von Arzneimitteln, in: Jusletter 2. September 2019.

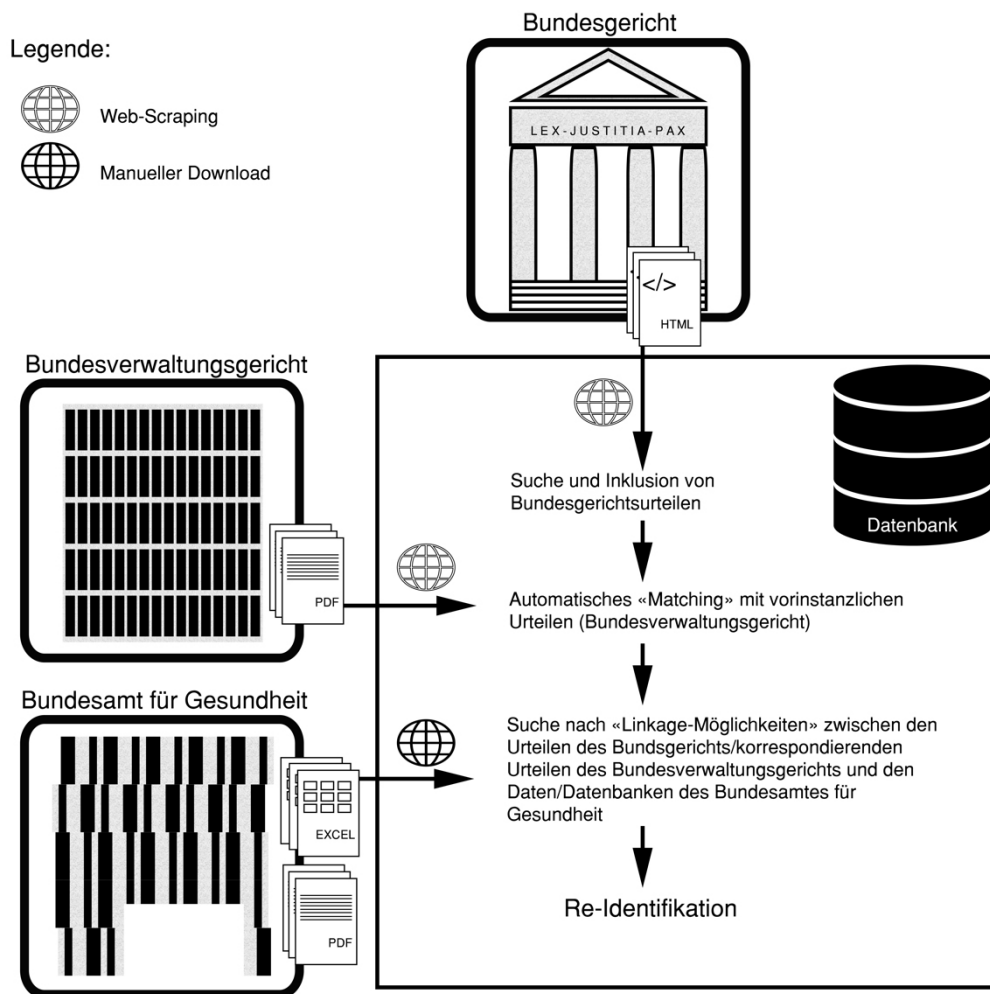


Abbildung 1: Überblick über das methodische Vorgehen und die Verbindung bzw. das «Linkage» von Daten(banken). CC BY SA 4.0-Lizenz.

[20] Gerade im Gesundheitsbereich ist davon auszugehen, dass eine Re-Identifikation noch einfacher möglich ist als in den anderen Lebensbereichen. Dies nur schon deshalb, weil die Fortschritte in der Bioinformatik das Sammeln zahlreicher genetischen Daten ermöglicht und es sich dabei um Daten handelt, die individueller nicht sein könnten und deshalb eine spezifische Re-Identifikation im Vergleich zu anderen Daten einfacher ermöglicht. Ersichtlich wird dies auch aus einer Studie aus dem Jahr 2013, die darlegte, dass die Gen-Sequenzierung in Kombination mit wenigen anderen Daten (Wohnsitzstaat in den USA, Alter) eine Re-Identifikation ermöglicht.²²

²² MELISSA GYMREK/AMY L. MCGUIRE/DAVID GOLAN/ERAN HALPERIN/YANIV ERLICH, Identifying Personal Genomes by Surname Inference, Science 2013;339:321–324.

[21] Vor dem Hintergrund dieser Entwicklungen stellt sich gerade in Bezug auf Gesundheitsdaten unweigerlich die Frage, ob und inwiefern es überhaupt anonymisierte bzw. anonyme Daten noch gibt. Gerade die Bio- und Medizininformatiker bezweifeln die Existenz dieser Datenkategorien. Noch zu klären gilt es, wie mit dieser zunehmenden Möglichkeit der Re-Identifikation umzugehen ist.²³

3. Fazit

[22] Die technologischen und medizinischen Entwicklungen in den letzten Jahren führen dazu, dass sich auch das rechtliche Verständnis, welche Daten als «Gesundheitsdaten» zu qualifizieren sind, verändert hat. Während sich Gesundheitsdaten früher primär aus der Arzt-Patienten-Beziehung ergaben, werden sie heute vermehrt direkt vom Patienten bzw. Konsumenten erfasst (z.B. über Apps oder Social Media).

[23] Auch der Kreis der involvierten Akteure hat sich erweitert. Neben den klassischen Akteuren bearbeiten vermehrt auch dem Gesundheitswesen bis vor wenigen Jahren noch fachfremde Akteure (z.B. Google, Amazon, Apple etc.) Gesundheitsdaten. Die Gesundheitsdaten beschränken sich dabei nicht auf klinische Daten, sondern umfassen auch vermehrt (genetische) Informationen vor klinischer Manifestation. Aus den Daten selbst ist nicht immer ersichtlich, ob es sich um *Gesundheitsdaten* handelt. Relevant ist der jeweilige Kontext. Die gleichen Daten (z.B. Anzahl zurückgelegter Schritte pro Tag) können bei einer (gesunden) Person als Lifestyle-Daten und bei einem Patienten (etwa bei einem Diabetiker oder herzinsuffizienten Patienten) als Gesundheitsdaten qualifiziert werden.

[24] Die technologischen Entwicklungen führen auch dazu, dass insbesondere mit der Methodik des «Linkage» anonymisierte und anonyme Daten vermehrt eine Re-Identifikation erlauben. Damit handelt es sich bei vermeintlich anonymisierten Daten oder anonymen Daten faktisch um Personendaten. Gesundheitsdaten erlauben – aufgrund ihrer regelmässig stärkeren Individualisierung gegenüber anderen Datenkategorien (zu denken ist etwa an genetische Daten, die jedem Individuum eigen sind) – mit einem kleineren Aufwand eine Re-Identifikation als andere Datenkategorien (z.B. Mobilitätsdaten).

Prof. Dr. iur. et Dr. med. KERSTIN NOËLLE VOKINGER, LL.M., Rechtsanwältin, Assistenzprofessorin Universität Zürich.

²³ Vgl. erste Überlegungen in Bezug auf Gerichtsurteile in VOKINGER/MÜHLEMATTER (Fn. 21).