D. Demmler, D. Krupka, H. Federrath. (Hrsg.): INFORMATIK 2022,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022 533

Datenschutzkonforme Weitergabe von Versichertendaten aus dem Forschungsdatenzentrum

Berna Orak[1], Erik Krempel[2], Arno Appenzeller[3]

Abstract: Die Nachfrage nach der breiten Verfügbarkeit von medizinischen Daten zu Forschungszwecken nimmt stetig zu. Die enormen Datenmengen bieten vor allem für Big Data Verfahren ein großes Potential. In der Gesetzgebung soll dieser Bedarf durch ein Forschungsdatenzentrum, das im Digitale-Versorgung-Gesetz (DVG) geregelt wird, erfüllt werden. Hierbei stellen sich allerdings eine Reihe von Fragen bezüglich des Datenschutzes. So sollen die Daten zwar in pseudonymisierter oder anonymisierter Form vorliegen, allerdings kann nach wie vor ein Re-Identifizierungsrisiko bestehen. Diese Ausarbeitung analysierte die bestehende Gesetzeslage und zieht überblickartig Vergleiche zu internationalen Vorschriften bezüglich der Regulierung anonymer Daten. Auf Basis dieser Analyse wird eine Erweiterung des Forschungsdatenzentrums skizziert, das mithilfe von Privatsphäre-wahrenden Technologien eine datenschutzkonforme Weitergabe von Versichertendaten ermöglichen kann.

Keywords: Forschungsdatenzentrum; Medizinischer Datenschutz; Datenschutzgrundverordnung; Medizinische Daten; Re-Identifizierung; Privatsphäre-wahrenden Technologien; Digitale-Versorgung-Gesetz

1 Einleitung

Das Digitale-Versorgung-Gesetz (DVG)[4] sieht in den §§ 303a-303f SGB V die Etablierung eines Forschungsdatenzentrums vor. Es soll den Zugang von Forscher:innen zu Daten von Millionen Krankenversicherten erleichtern, um unter anderem eine bessere Nutzbarkeit von Gesundheitsdaten für Forschungszwecke zu ermöglichen.[5] Neu ist, dass die Versorgungsdaten von ca. 70 Millionen Bürger:innen zentral gesammelt und der Forschung zur Verfügung gestellt werden. Der Umfang der gesammelten Daten in Verbindung mit ihrer besonderen Vulnerabilität ist deutschlandweit einzigartig und hat in Politik sowie

[1] Goethe-Universität Frankfurt am Main, Lehrstuhl für Öffentliches Recht, Informationsrecht, Umweltrecht und Verwaltungswissenschaft, Institut für Öffentliches Recht, Theodor-W.-Adorno-Platz 4, 60629 Frankfurt am Main, Deutschland, orak@jur.uni-frankfurt.de
[2] Fraunhofer Institut für Optronik, Systemtechnik und Bildverarbeitung (IOSB), Fraunhoferstr. 1, 76131 Karlsruhe, Deutschland, erik.krempel@iosb.fraunhofer.de
[3] Karlsruher Institut für Technologie, Lehrstuhl für Interaktive Echtzeitsysteme, c/o Technologiefabrik Haid-und-Neu-Str. 7,76131 Karlsruhe, Deutschland arno.appenzeller@kit.edu
[4] Gesetz vom 09.12.2019, BGBl. I S.2562.
[5] BT-Drucksache, 19/13438, S. 2.

doi:10.18420/inf2022_45

Wissenschaft Aufmerksamkeit erregt und ebenso Kritik hervorgerufen[6], obwohl die Daten zum Teil anonym zur Verfügung gestellt werden sollen.

Im Rahmen von Forschungsdatenverarbeitung kristallisiert sich längst heraus, dass eine Re-Identifizierung von Daten nicht mehr auszuschließen ist.[7] Gerade im Gesundheitssektor wird die Verarbeitung selbst von anonymen Daten als gefahrvoll betrachtet, da aufgrund der Einzigartigkeit der Angaben technische Mechanismen oftmals keinen allumfassenden Schutz vor einer De-Anonymisierung gewährleisten können.[8] Dabei beschränkt sich die Schwierigkeit nicht nur auf den Gesundheitssektor.[9]. Besonders riskant ist die Weitergabe der Daten an Verarbeiter, die bereits über unterschiedliches Zusatzwissen aus anderen Quellen verfügen. Sie würden zudem bei langfristiger Aufbewahrung von zukünftigen technischen Entwicklungen profitieren, weil sie dann vorhandene Daten genauer auswerten und möglicherweise einzelnen Personen zuordnen könnten.[10] Eine Re-Identifizierungen ist dann nicht ausschließbar.[11]

Behält man diesen Umstand im Blick, stellen sich mit der Etablierung eines Forschungsdatenzentrums, das die Weitergabe der Versorgungsdaten von ca. 90% der Bevölkerung, nämlich der gesetzlich Krankenversicherten, an unbestimmte Nutzungsberechtigte erlaubt, viele Fragen. Wer hat Zugriff auf die Daten? Wie sieht die technische Umsetzung aus? Wie wird verhindert, dass eine solche Re-Identifikation doch möglich ist? Das nach den §§ 303a-303f Sozialgesetzbuch Fünftes Buch (SGB V) vorgesehene Forschungsdatenzentrum lässt diese und weitere Fragen nach der technischen Umsetzung einer Anonymisierung sowie Pseudonymisierung offen. Auch die DSGVO greift die Anonymisierung nicht ausdrücklich auf. Blickt man in die USA und Japan, bestehen dagegen Rahmenregelungen zur Vermeidung eines Re-Identifikationsrisikos.

Der folgende Beitrag betrachtet technische Ausgestaltungsmöglichkeiten zur datenschutzkonformen Weitergabe von Daten an Dritte mit Bezug zum Gesundheitswesen. Im Anschluss an einen Überblick über die Funktionsweise des Forschungsdatenzentrums erfolgt eine Auseinandersetzung mit den rechtlichen Rahmenbedingungen einer Anonymisierung (2). Als entscheidende Vorfrage werden die Herausforderungen der Datenweitergabe an Dritte aus technischer Sicht erörtert (4), um dann der Frage nach Privatsphäre wahrenden Technologien zur datenschutzkonformen Datenweitergabe nachzugehen (5). Ein Fazit schließt den Beitrag (6).

[6] Siehe [Bu19], [SgDB20], [BSgD20], [We20],[KS20].
[7] [MH20], 3574, [RG21], 487.
[8] Caspar, [SHD19], Art. 89 Rn. 55; [La16], 363.
[9] Siehe dazu [RG21], 487.
[10] [Ro21], 176.
[11] [Ro21], 176.

2 Die Datenaufbereitung im neuen Forschungsdatenzentrum und ihre Gefahrpotenziale

Das Forschungsdatenzentrum stellt eine Weiterentwicklung der Datenaufbereitungsstelle dar, die seit 2013 Datenanalysen mit Abrechnungsdaten aus der gesetzlichen Krankenversicherung zu Forschungszwecken durchführt.[12] Es werden Angaben zum Alter, Geschlecht, Wohnort und ICD-Codes im Forschungsdatenzentrum gesammelt.[13] Der ICD-Code ist eine standardisierte Codierung von medizinischen Diagnosen, die von jedermann mithilfe einer Tabelle aufgelöst werden kann. Das bedeutet: Unabhängig davon, ob die Daten pseudonymisiert oder anonymisiert vorliegen, der Datensatz enthält zunächst hochsensible Angaben.

Vorgesehen ist, dass die Datenbereitstellung von drei öffentlichen Stellen des Bundes übernommen wird, vgl. § 303a Abs. 1 SGB V. Der Spitzenverband Bund der Krankenkassen wirkt als sog. Datensammelstelle und das Robert-Koch-Institut (RKI) als sog. Vertrauensstelle. Die Aufgaben des Forschungsdatenzentrums nimmt das Bundesinstitut für Arzneimittel und Medizinprodukte wahr.[14] Die gesetzlichen Krankenkassen geben an die Datensammelstelle, die Angaben für jeden Versicherten mit einem Lieferpseudonym weiter, vgl. § 303b SGB V. Nach Überprüfung auf ihre Vollständigkeit werden die Daten mit einer Arbeitsnummer gekennzeichnet und ohne Lieferpseudonym an das Forschungsdatenzentrum weitergegeben. Das Lieferpseudonym und die Arbeitsnummer erhält nur die Vertrauensstelle, also das RKI. Das Forschungsdatenzentrum übernimmt sodann, wenn Daten zur Forschung angefragt werden, die Aufbereitung der beantragten Daten sowie die Qualitätssicherung und leitet sie an die nutzungsberechtigen Stellen nach Überprüfung der Anträge weiter, vgl. § 303d Abs. 1 SGB V. Die Nutzungsberechtigen erhalten die Daten grundsätzlich in anonymisierter und aggregierter Form.[15] Darüber hinaus prüft das Forschungsdatenzentrum ein spezifisches Re-Identifikationsrisiko beim jeweiligen Nutzungsberechtigen.[16] Die Anforderungen an die technische Umsetzung einer Minimierung des Re-Identifikationsrisikos bleiben dem Forschungsdatenzentrum bzw. dem jeweiligen Auftragsverarbeiter überlassen.[17] Es besteht immerhin die Möglichkeit, die Datenbereitstellung mit der Auflage zu verbinden, eine Zusammenführung der Daten mit externen Datenbeständen zu unterlassen.[18] Insgesamt entspricht die Aufgabenwahrnehmung des Forschungsdatenzentrums somit der Art nach einem Datentreuhänder[19]. Eine allgemeingültige Definition eines Datentreuhandmodells

[12] Michels, [BK22], § 303a Rn. 1a.
[13] Siehe zum vollzähligen Umfang der weitergegebenen Daten: § 303b Abs. 1 i.V.m. § 3 Abs. 1 DaTraV.
[14] Siehe auch § 1 DaTraV.
[15] In Ausnahmefällen können die Daten auch in pseudonymisierter Form zur Verfügung gestellt werden, siehe § 303e Abs. 4 Satz 1 SGB V.
[16] Vgl. § 303d Abs. 1 Nr. 5 SGB V.
[17] Siehe § 303b Abs. 1 Satz 2 und § 303b Abs. 3 Satz 3.
[18] Siehe § 8 Abs. 3 DaTraV.
[19] Siehe [Bu21], 806 ff.

gibt es nicht.[20] Grundsätzlich stellt ein Datentreuhänder einen neutralen Vermittler zwischen Datenverarbeiter und betroffener Person dar, der keinen Mehrwert aus den Daten schöpft. [21]

Die Schwierigkeit der Kontrolle über die Daten beginnt mit der Weitergabe der Daten an die Nutzungsberechtigten. Sie ist gesteigert, wenn sie mit anderen Datensätzen zusammengeführt werden. Die gespeicherten Gesundheitsdaten können z.B. mithilfe von Big Data Analysen ausgewertet werden. Charakteristisch für Big Data Anwendungen ist, dass große Datenmengen (Volume), unterschiedliche Datentypen- und quellen (Variety) sowie eine hohe Geschwindigkeit der Generierung und Verarbeitung der Daten (Velocity) vorliegen.[22] Durch die massenhafte Auswertung unterschiedlicher Daten und Datenquellen können Verknüpfungen hergestellt werden, die wiederum die Identität des Datensubjekts enthüllen oder es jedenfalls identifizierbar machen.[23] Dagegen war die Verarbeitung anonymer Daten in der Vergangenheit eine gut beherrschte Technik, um Identifizierungen zu vermeiden.[24] Dies folgt auch daraus, dass vor wenigen Jahren die Verknüpfung von Datenbanken nur mit unverhältnismäßigem Aufwand möglich war. Mit der Einführung eines Forschungsdatenzentrums werden nun große Datenmengen mit hoher Qualität und Aussagekraft zugänglich.

Zum aktuellen Zeitpunkt befindet sich das Forschungsdatenzentrum noch im Aufbau. Es werden die rechtlichen, technischen personellen und organisatorischen Maßnahmen des neuen Forschungsdatenzentrums definiert und implementiert.[25]. Daneben ist vor dem Bundesverfassungsgericht eine Verfassungsbeschwerde gegen Teile des Digitale-Versorgung-Gesetzes anhängig, so dass noch nicht klar ist, inwieweit es tatsächlich eingeführt wird. Eine einstweilige Anordnung gegen das Gesetz mit dem Ziel Teile des DVG bis zum Hauptsacheverfahren außer Kraft zu setzen, war allerdings erfolglos, da nach Auffassung des Bundesverfassungsgerichts mit der Außerkraftsetzung nachteilige Folgen für wichtige Gemeinwohlbelange verhindert würden (u.a. die medizinische Forschung).[26] Die Vorwürfe gegen das Gesetz sind aber nicht haltlos. Die Diskussion ist geprägt vom fehlenden Widerspruchsrecht der Bürger:innen gegen die Verarbeitung der Daten[27], von der Streubreite und Massenhaftigkeit des Eingriffs bis hin zu europarechtlichen Problemen.[28] Aber eben auch und besonders ein potenzielles Re-Identifikationsrisiko durch die Verknüpfung mit eigenen

[20] Siehe hierzu [K1], 784 f.; [Bu21], 807.
[21] [K1], 784 f.; [Bu21], 807 f.
[22] [De17], 54.
[23] [RG21], 487.
[24] Schaar, [St17], S. 144.
[25] Bundesinstitut für Arzneimittel und Medizinprodukte, https://www.forschungsdatenzentrum-gesundheit.de/das-fdz (zuletzt aufgerufen am 08.07.2022)
[26] BVerfG, Beschluss vom 19. März 2020- 1 BvQ 1/20; siehe hierzu auch: [SgDB20]; von Dewitz, [Ro18], § 303a, Rn. 2.
[27] [We20], 542 f.; anders: Bitkom, Pressemitteilung v. 07.Novemeber 2019, https://www.bitkom.org/Presse/Presseinformation/Bitkom-zum-Digitale-Versorgung-Gesetz (zuletzt aufgerufen am 11.04.2022).
[28] Siehe [SgDB20]; [BSgD20], 990; zum Ausschluss der Betroffenenrechte siehe: [We20], 542 f.; zur Abgrenzung von pseudonymisierten und anonymisierten Daten: [KS20], 43 ff.

Datenbanken der Nutzungsberechtigen ist trotz anonymisierter Datensätze problematisch und hat u.a. den Beschwerdeführer[29] zu seiner Verfassungsbeschwerde bewegt.

3 Rechtliches Fundament anonymer Daten

Anknüpfend an den bereits dargelegten Aufbau des Forschungsdatenzentrums nach dem DVG werden nun die rechtlichen Anforderungen an eine Anonymisierung nach der Datenschutz-Grundverordnung[30] (DSGVO) untersucht sowie ein Überblick über den Umgang Japans und der USA mit anonymen Daten und dem Re-Anonymisierungsrisiko gegeben.

3.1 Art. 9 Abs. 1 und Abs. 2 DSGVO

Art. 4 Nr. 15 DSGVO[31] legaldefiniert als besondere Daten „personenbezogene Daten, die sich auf die körperliche oder geistige Gesundheit einer natürlichen Person, einschließlich der Erbringung von Gesundheitsdienstleistungen, beziehen und aus denen Informationen über deren Gesundheitszustand hervorgehen". Die im Forschungsdatenzentrum gespeicherten Daten stellen solche Gesundheitsdaten dar. Die Verarbeitung von Gesundheitsdaten ist grundsätzlich verboten, vgl. Art. 9 Abs. 1 DSGVO, und nur ausnahmsweise erlaubt. Sie geht mit erheblichen Gefahren für die Grundfreiheiten und Grundrechte der betroffenen Person einher.[32] Der Gesetzgeber hat daher in Art. 9 Abs. 2 DSGVO besondere Voraussetzungen für die ausnahmsweise Verarbeitung von sensiblen Daten geschaffen. Diese Öffnungsklauseln erlauben den Mitgliedstaaten weitergehende mitgliedstaatliche Regelungen zu erlassen.[33] Neben der Einhaltung der allgemeinen Prinzipien aus Art. 5 DSGVO, wie beispielsweise die Zweckbindung, Datensparsamkeit und Transparenz sowie einer Rechtsgrundlage, Art. 6 Abs. 1 DSGVO[34], stellt Art. 9 Abs. 2 DSGVO darüberhinausgehende spezifische Anforderungen für die Verarbeitung von Gesundheitsdaten auf.

Die Weitergabe der Versorgungsdaten durch die Krankenkassen an das Forschungsdatenzentrum auf der Basis von §§ 303a-f SGB V wird durch die Öffnungsklauseln des Art. 9 Abs. 2 lit. h und lit. j) DSGVO ermöglicht. Dies setzt u.a. zusätzlich voraus, dass die Verarbeitung für wissenschaftliche Forschungszwecke angemessene und spezifische Maßnahmen zur Wahrung der Grundrechte der betroffenen Person vorsehen und die Verarbeitung gem. Art. 89 Abs. 1 DSGVO erforderlich ist.

[29] Der Beschwerdeführer leidet an einer seltenen Erbkrankheit und befürchtet trotz Pseudo- oder Anonymisierung aus den Datensätzen re-identifiziert werden zu können.

[30] VO (EU) 2016/679.

[31] Die Kompetenz der EU für das Datenschutzrecht ergibt sich aus Art. 16 Abs. 2 UAbs. 1 S. 1 AEUV, siehe auch Hornung/Spiecker gen. Döhmann, [SHD19], Einleitung, Rn. 155 ff.

[32] Petri, [SHD19], Art. 9, Rn. 1.

[33] Petri, [SHD19], Art. 9, Rn. 101 f.

[34] Zum Verhältnis zwischen Art. 9 und Art. 6 DSGVO: Wedde, [D0], Art. 9, Rn. 3; Petri, [SHD19], Art. 9, Rn. 3; Weichert, [KB20], Art. 9 Rn. 4,; aA: Schantz, [SW17], Rn. 705.

3.2 Art. 89 DSGVO

Art. 89 stellt keinen Erlaubnistatbestand dar, sondern knüpft die an sich zulässige Verarbeitung nach Art. 6 DSGVO und Art. 9 DSGVO nochmals an bestimmte Anforderungen.[35] Damit bezweckt die Vorschrift u.a., den Datenminimierungsgrundsatz aus Art. 5 Abs. 1 lit. c DSGVO zu gewährleisten.[36] Darüber hinaus dient die Regelung dazu, einen Ausgleich zwischen der informationellen Selbstbestimmung und den in der DSGVO gewährten Privilegierungen zugunsten der Forschung zur Berücksichtigung der Forschungsfreiheit herzustellen.[37]. Der Vorbehalt geeigneter Garantien für die Rechte und Freiheiten der betroffenen Person erfordert, „dass technische und organisatorische Maßnahmen bestehen, mit denen insbesondere die Achtung des Grundsatzes der Datenminimierung gewährleistet wird“. Die Maßnahmen in Art 89 Abs. 1 DSGVO verfolgen einen abgestuften Ansatz.[38] Sofern die Zwecke der Verarbeitung mittels einer Anonymisierung erreicht werden können, sind die Daten zu anonymisieren.[39] Scheidet dagegen eine Anonymisierung aus, wenn die Zwecke der Verarbeitung sich nur mit personenbezogenen Daten erreichen lassen,[40] ist zunächst eine Pseudonymisierung in Betracht zu ziehen. Erst wenn dies nicht möglich ist, dürfen unmittelbar erkennbare personenbezogene Daten verarbeitet werden.

3.3 Anonymisierung und die DSGVO

Die Vorprüfung der Möglichkeit einer Anonymisierung wird in Art. 89 DSGVO, im Gegensatz zur Pseudonymisierung, nicht ausdrücklich erwähnt. Auch sonst wird der Begriff nicht in den Artikeln der DSGVO verwendet. Lediglich in Erwägungsgrund (EG) 26 zur DSGVO finden sich Erläuterungen zur „Anonymisierung“. Unter anonymen Informationen werden nur solche Daten verstanden, „die sich nicht auf eine identifizierte oder identifizierbare natürliche Person beziehen, oder personenbezogene Daten, die in einer Weise anonymisiert worden sind, dass die betroffene Person nicht oder nicht mehr identifiziert werden kann“.[41] Daraus lässt sich schließen, dass anonyme Daten das Gegenteil von personenbezogenen Daten sind.[42] Die Grundsätze des Datenschutzes gelten nicht für anonyme Informationen, vgl. EG 26 S. 5 zur DSGVO. Somit ist der sachliche Anwendungsbereich der DSGVO bei anonymen Daten nicht eröffnet. Die zitierte Definition findet sich aber “nur“ in einem Erwägungsgrund. Diese entfalten keine rechtliche Bindungswirkung.[43] Ihnen kommt nur

[35] Pauly, [PP21], Art. 89, Rn. 1; Buchner/Tinnefeld, [KB20], Art. 89 Rn. 1.
[36] Buchner/Tinnefeld, [KB20], Art. 89 Rn. 19.
[37] [Ro19], 159.
[38] Pauly, [PP21], Art. 89, Rn. 12.
[39] Caspar, in [SHD19], Art. 89, Rn. 52.
[40] Pauly, in: [PP21], Art. 89, Rn. 12.
[41] Siehe EG 26 S. 5 zur DSGVO.
[42] Karg, [SHD19], Art. 4 Nr. 1. Rn. 20; [RG21], 487; [Gi21], 482.
[43] Zur rechtlichen Qualität von Erwägungsgründen von Gemeinschaftsrechtsakten: EuGH, Urt. v. 19.6.2014, Rs. C-345/13, ECLI:EU:C:2014:2013 – Karen Millen Fashions, Rn. 31.

besondere Beachtung bei der Auslegung der Vorschriften der DSGVO zu.[44] Da bisher keine Definition der Anonymisierung in der DSGVO vorzufinden ist, bietet es sich an, die Terminologie nach der alten Rechtsauffassung des BDSG a.F. heranzuziehen und zu analysieren.

Ursprünglich wurde das Anonymisieren als „das Verändern personenbezogener Daten derart, dass die Einzelangaben über persönliche oder sachliche Verhältnisse nicht mehr oder nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmbaren natürlichen Person zugeordnet werden können“ definiert, vgl. § 3 Abs. 6 BDSG a.F. Bei der Beurteilung, welche Kriterien für den Begriff anonymer Daten im heutigen Recht heranzuziehen sind, kann die nationale Rechtslage a.F. allerdings nur begrenzt Orientierung bieten. Ein direkter Rückgriff wäre ohnehin unionsrechtswidrig.[45] Es sind nicht mehr nur der „unverhältnismäßig hohe Aufwand an Zeit, Kosten und Arbeitskraft“ zu berücksichtigen, sondern auch andere Faktoren, die für eine Identifizierbarkeit sprechen.[46]

Nach heutigem Verständnis der DSGVO liegen anonymisierte Daten vor, wenn kein Personenbezug möglich ist.[47] Dabei muss mit hinreichender Wahrscheinlichkeit feststehen, dass eine Re-Identifizierung ausgeschlossen ist. [48] Das Entfernen des Namens, der Anschrift und des Geburtsdatums genügen hierfür nicht.[49] Die Bewertung der Wahrscheinlichkeit des Risikoeintritts erfordert auch, dass der der Verantwortliche einen Personenbezug durch Dritte zu vertreten hat, selbst wenn er aus ex-ante Perspektive diesen nicht herstellen konnte.[50] Nach EG 26 „sollten alle objektiven Faktoren, wie die Kosten der Identifizierung und der dafür erforderliche Zeitaufwand, herangezogen werden, wobei die zum Zeitpunkt der Verarbeitung verfügbare Technologie und technologische Entwicklungen zu berücksichtigen sind.“ Der EuGH hat dies präzisiert, dass es vorrangig auf die Mittel ankommt, die legal herangezogen werden können.[51] Damit wird zum Ausdruck gebracht, dass das Datenschutzrecht kein erlaubtes Risiko zulässt.[52] Wie unterschiedlich die Anforderungen über Zeitablauf werden können, zeigt sich schon daran, dass mit den heutigen technischen Möglichkeiten und mit dem Abgleich von Datenbanken an das Merkmal ”unverhältnismäßig hoher Aufwand” durchaus andere Anforderungen zu stellen sind als noch zu BDSG a.F.-Zeiten.[53] Erforderte die Verknüpfung mit anderen Datenbanken früher noch einen hohen zeitlichen Aufwand, ist das mit den heutigen technischen Möglichkeiten in wenigen Sekunden erledigt.

Festzuhalten bleibt, dass die Definition anonymer Daten ein changierender ist, der unter

[44] Wegener, [CR21], Art. 19 Rn. 32.
[45] [Me21], der auch die landesrechtlichen Legaldefinitionen in § 3 BbgDSG, § 4 DSG NRW und § 28 Abs. 3 ThürDSG für unionswidrig einstuft; dagegen: [RG21], 487 , die die Definition weiterhin aufgreifen.
[46] Hierfür spricht der Wortlaut ”wie” im EG 26 S. 4 zur DSGVO; siehe: [Me21]; Ernst, [PP21], Art. 4 Rn. 50.
[47] [Ro19], 159; [MH20], 3574.
[48] [MH20], 3574.
[49] [MH20], 3574.
[50] Ernst [PP21], Art. 4 Rn. 13.
[51] EuGH, Urt. v. 19.10.2016, Rs. C-582/14, – Breyer.
[52] [MH20], 3574; Ernst, [PP21], Art. 4 Rn. 13.
[53] Ernst, [PP21], Art. 4 Rn. 50.

dem technischen Wandel stetig einer Neujustierung bedarf. Dies ruft für den Umgang mit anonymen Daten nach einer Rahmenregelung, die bislang noch nicht existiert. Umso erfreulicher ist das Vorhaben der Bundesregierung, Standards für die Anonymisierung und die Sanktionierung einer De-Anonymisierung etablieren zu wollen.[54] Ob die Etablierung dieser Regelungen ausschließlich auf nationaler Ebene erfolgt oder eine europäische Lösung bestrebt wird, bleibt unklar. Angesichts der europaweiten Wirkung der DSGVO wäre eine europäische Regulierung des Umgangs mit anonymisierten Daten wünschenswert, auch im Hinblick auf den gerade entstehenden Datenmarkt in der EU durch Gesetzeswerke wie z.B. den Data Governance Act, jedenfalls aber für besonders sensible Daten wie anonymisierte Gesundheitsdaten, die auf supranationaler Ebene im Europäischen Gesundheitsdatenraum ausgetauscht werden sollen.[55]

3.4 Überblick USA und Japan

In den USA existiert mit dem überarbeiteten Health Insurance Portabilitity and Accountability Act (HIPAA)[56] seit 2013 eine Regelung für die standardisierte Übermittlung von elektronischen Gesundheitsdaten, die zumindest die Gefahr einer Re-Identifikation berücksichtigt.[57] Der HIPAA listet 18 Merkmale (Name, ortsbezogene Angaben, Sozialversicherungsnummer etc.), die zur Umsetzung einer De-Identifikation entfernt werden müssen.

In Japan dagegen ist der Gesetzgeber einen Schritt weitergegangen und hat die Definition anonymer Daten in den Amended Act on the Protection of Personal Information (APPI)[58] aufgenommen, vgl. Art. 2 Abs. 9 APPI. Im Zuge der Reform des APPI wurden auch Standardisierungsprozesse zu Anonymisierungstechniken, Anforderungen an die Datensicherheit, ein Verbot der Zusammenführung unterschiedlicher Datenquellen sowie Informationspflichten gegenüber der Öffentlichkeit eingeführt.[59]

Der aufgezeigte kursorische Überblick illustriert schon auf den ersten Blick Unterschiede in der Regulierung anonymer Daten, die zwischen Europa, Amerika und Asien herrschen. Unabhängig davon, wie effektiv die ergriffenen Maßnahmen anderer Rechtsordnungen sind, gilt es die Regelungslücke für die Anwendung anonymisierter Daten und damit die Nicht-Anwendbarkeit der DSGVO zu schließen. Die Öffnungsklauseln des Art. 9 Abs. 2 DSGVO ermöglichen dies zumindest für Gesundheitsdaten und damit auch für das Forschungsdatenzentrum. Die Herausforderung für die Umsetzung des Forschungsdatenzentrums liegt

[54] Koalitionsvertrag der Bundesregierung, S.17, https://www.tagesschau.de/koalitionsvertrag-147.pdf (zuletzt aufgerufen am 08.07.2022).
[55] Siehe zum Europäischen Gesundheitsdatenraum: COM (2022) 197 final.
[56] Pub. L 104-191, 104th Congress, 110 Stat.1936 v. 21.8.1996; 42 U.S.C. 210ff, 45 C.F.R. part 160 and subparts A und E of part 164.
[57] [Le13], 767.
[58] Gesetz Nr. 57 von 2003; in Kraft getreten am 1.4.2005.
[59] Siehe hierzu: [GLF18], 417f

aber nicht nur am schillernden Anonymisierungsbegriff, sondern auch in der Auswahl der einschlägigen technischen Infrastruktur, um ein minimales Re-Identifikationsrisiko zu gewährleisten.

4 Herausforderungen der Datenweitergabe an Dritte

Aus einer rein datenschutzrechtlichen Betrachtung wäre es für den Schutz der Rechte und Freiheiten der von der Datenübertragung betroffenen Personen vorteilhaft, wenn diese in einer Form erfolgt, die der Definition von anonymen Daten unter BDSG a. f. entspricht. Das würde bedeuten, dass der Personenbezug von Daten derart aufgehoben ist, dass er nicht oder nur unter unverhältnismäßigem Aufwand an Zeit, Kosten und Arbeitskräften wiederhergestellt werden kann.

Dieser Wunsch ist jedoch in der Realität nur selten erreichbar. Menschen sind nicht nur durch eindeutige Identifikatoren wie Name oder Krankenversichertennummer zu identifizieren. Ebenso lassen sich Personen über eine Kombination uneindeutiger Daten, oft Quasi-Identifikatoren genannt, wie Geschlecht, Geburtsdatums oder die Postleitzahl identifizieren. Besonders eindrücklich demonstrierte dies Latanya Sweeney im Jahr 2000.[60] In ihrer Veröffentlichung hat sie gezeigt, dass sich 87% aller Amerikaner eindeutig aus der Kombination des Geburtsdatums, ihres Geschlechts und ihrer Postleitzahl identifizieren lassen. Da medizinische Versorgungdaten noch deutlich detaillierter sind, muss man davon ausgehen, dass damit ebenfalls eine Re-Identifizierung möglich ist.

Cynthia Dwork geht hier sogar so weit, dass sie sagt: “De-Identified data isn’t”. Damit möchte sie auf die besondere Abhängigkeit zwischen der Anonymisierung von Daten und deren Nützlichkeit eingehen. Wenn garantiert werden soll, dass Angreifer nicht in der Lage sind Personen zu Re-Identifizieren, müssen die Daten sehr stark verändert werden. Das bedeutet, dass alle Identifikatoren entfernt und auch alle Quasi-Identifikatoren so stark verändert werden müssen, dass über diese keine Zuordnung mehr möglich ist. Das kann beispielsweise erreicht werden, in dem das Geburtsdatum auf Monat und Jahr reduziert wird oder in der Postleitzahl die hinteren Stellen entfernt werden. Offensichtlich ist, dass dadurch ein Teil der Aussagekraft der Daten verloren geht. Je mehr die Daten verändert werden, um eine Anonymisierung zu erreichen, desto stärker wird die Nützlichkeit eingeschränkt. Laut Dwork kann nur entschieden werden, ob etwas erzeugt werden soll, was sicher anonymisiert ist oder ob nützliche Daten zurückbehalten werden sollen.

An dieser Stelle wird deutlich, dass der Zweckbindungsgrundsatz in Wechselwirkung mit dem Prinzip der Datensparsamkeit steht[61]. Die Daten dürfen nur für einen bestimmten Zweck verarbeitet werden; gleichzeitig sind sie damit auf das notwendige Maß für diesen Zweck zu beschränken. Der Grundsatz der Datensparsamkeit darf aber nicht dazu führen, dass unbrauchbare Daten erzeugt werden, die für das Forschungsvorhaben letztlich keine

[60] [Sw00]
[61] Siehe zu den Prinzipien der DSGVO Art. 5 Abs. 1 DSGVO.

Aussagekraft haben. Tatsächlich verlangt das Prinzip der Datenminimierung auch nicht nach einer absoluten Beschränkung oder Reduzierung der Datenmenge.[62] Um die Nützlichkeit der Daten zu gewährleisten, muss daher auf die Festlegung eines bestimmten Zwecks geachtet werden. Ist im Ergebnis der Verarbeitungszweck durch anonymisierte Daten erreichbar, würde eine Zuwiderhandlung durch Verarbeitung nicht anonymisierter Daten gegen das Prinzip der Datenminimierung verstoßen.[63] Für die Forschung mit und an Daten stellt sich aber häufiger die Hürde, von Beginn an einen bestimmten Zweck festzulegen.[64] Darüber hinaus steht der Anonymisierung gerade im klinischen Forschungsbereich entgegen, dass eine fortlaufende Zuordnung der Daten bei Langzeitstudien notwendig ist[65].

In der medizinischen Forschung wird oft davon ausgegangen, dass eine echte Anonymisierung bei Erhalt der Nützlichkeit der Daten nicht möglich ist. Gleichzeitig sollen alle möglichen technischen und organisatorischen Maßnahmen ergriffen werden, um die Daten bzw. die betroffenen Personen optimal zu schützen. Die TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. – erarbeitet seit ihrer Gründung im Jahr 2003 Konzepte für eine datenschutzgerechte medizinische Forschung. Ein zentraler Baustein vieler Konzepte ist die Trennung der Daten in die zwei Gruppen IDAT und MDAT. IDAT sind dabei alle Daten wie Namen, Adresse oder Krankenversichertennummer die primär für die Identifikation der Personen benötigt werden. MDAT bezieht sich auf alle Daten die Aussagen über medizinische Sachverhalte wie beispielsweise ein Laborergebnis oder einer Diagnose machen. IDAT werden im Rahmen der Konzepte so verändert, dass ein optimaler Schutz der betroffenen Personen erreicht wird. Beispielsweise indem sie komplett durch ein Pseudonym ersetzt werden und nur noch eine Treuhandstelle eine Zuordnung zwischen Pseudonym und IDAT herstellen kann. MDAT bleiben unverändert und werden beispielsweise dadurch geschützt, dass strenge Zugangs- und Zugriffsregelungen etabliert werden. Das Ziel ist Datensparsamkeit zu gewährleisten, damit Wissenschaftler:innen nur genau die Daten bekommen, die sie für ihre Forschung benötigen. Aufbauend auf den Konzepten der TMF können nun noch weitere Maßnahmen ergriffen werden, um die betroffenen Personen bei einer Weitergabe der Daten optimal zu schützen.

5 Technisches Konzept für ein datenschutzkonformes Forschungsdatenzentrum

Der folgende Abschnitt gibt einen Überblick über verfügbare und nutzbare Privatsphäre wahrende Technologien zur Datenweitergabe. Auf Basis der rechtlichen Analyse aus den vorherigen Abschnitten und den hier aufgezeigten Technologien wird ein Umsetzungsvorschlag zu einem datenschutzkonformen Forschungsdatenzentrum aufgezeigt.

[62] Herbst, [KB20], Art. 5 Rn. 5; [K0] 185.
[63] Herbst, [KB20], Art. 5 Rn. 56.
[64] Roßnagel, [SHD19], Art. 5 Rn. 128; siehe auch: [Sp17].
[65] Weichert, ABIDA Gutachten, 2018, Big Data im Gesundheitsbereich, S. 192, https://www.abida.de/sites/default/files/ABIDA%20Gutachten-Gesundheitsbereich.pdf.

5.1 Privatsphäre wahrende Technologien zur datenschutzkonformen Datenweitergabe

In der Vergangenheit haben sich vor allem für medizinische Studien mit Datenveröffentlichung /-weitergabe Metriken wie k-Anonymität und l-Diversity etabliert.[66] Bei k-Anonymität wird für eine Datenveröffentlichung vorausgesetzt, dass es für jeden Eintrag mit einem als sensitiv erachtetes Attribut (beispielsweise ein medizinischer Befund) $k-1$ andere Datensätze mit demselben sensitiven Attribut existieren.[67] Hierdurch soll das triviale Raten mit Hintergrundwissen über eine seltene Erkrankung verhindert werden. Gruppen mit denselben Attributen werden Äquivalenzklassen genannt. Diese Klassen erhält man durch die Unterdrückung oder Generalisierung von identifizierenden Attributen. Während bei der Unterdrückung das einfache Entfernen beziehungsweise Ersetzen durch einen Platzhalter ausreicht, werden für das Generalisieren Attributs Hierarchien benötigt. So wird zum Beispiel ein spezifisches Alter einer Altersgruppe zugeordnet. Allerdings besitzt k-Anonyimität diverse Schwächen. Sollte es beispielsweise zu einer unabsichtlichen Veröffentlichung der Daten kommen, ist es für die betroffenen Personen irrelevant, dass auch weitere Personen das gleiche sensitive Attribut haben. Um solche Angriffe zu verhindern, wurde die Definition mit l-diversity verschärft.[68] Um die Äquivalenzklassen weiter zu diversifizieren und einen Angriff mit Hintergrundwissen zu erschweren, wird gefordert, dass jede Äquivalenzklasse mindestens l verschiedene sensitive Attribute beinhaltet. Hierdurch würde das einfache Wissen über die Zugehörigkeit eines Individuums zu einer Äquivalenzklasse nicht ausreichen, um das sensitive Attribut zu erhalten. Allerdings gibt es auch eine Reihe effizienter Hintergrundwissens Angriffe auf l-diversity. Beispielsweise die Ähnlichkeitsattacke, bei der die sensitiven Attribute sehr ähnlich oder semantisch synonym sind (beispielsweise verschiedene Arten von Darmkrebs). Neben l-diversity existieren weitere Begriffe wie t-closeness und δ-presence , die aber ebenfalls Angriffsvektoren bieten und weitere nicht immer trivial erfüllbare Anforderungen an die zu weitergebende Daten stellen.[69]

In den letzten Jahren wurde der Begriff der Differential Privacy (DP) nach Dwork ein immer relevanteres Maß für eine Privatsphäre Garantie.[70] Im Gegensatz zu den vorher genannten Techniken, benötigt Differential Privacy weniger semantisches Wissen über die zu privatisierenden Daten. Vielmehr betrachtet DP, wie sehr die Daten eines Individuums das Ergebnis einer Datenauswertung beeinflusst. Gemäß der Definition von Differential Privacy bedeutet dies für zwei Datensätze D_1 und D_2, die sich um genau ein Element unterscheiden, und $S \subseteq Bild(K)$, dass gilt

$$\Pr[\mathcal{K}(D_1) \in S] \leq \exp(\epsilon) \times \Pr[\mathcal{K}(D_2) \in S].$$

Für eine gegebene Auswertung der Daten, bedeutetet dies, dass für alle möglichen Datensätze das Ergebnis sich höchstens um e^{ϵ} unterscheidet. Um diese Eigenschaft zu erfüllen wird K

[66] [Zu21]
[67] [Sw02]
[68] [Ma06]
[69] [LLV07, NAC07]
[70] [Dw06]

als sogenannter Differential Private Mechanismus genutzt. Mit ihm wird das Ergebnis so stark verändert, bis die Bedingung von DP erfüllt ist. ϵ wird auch das Privatsphärebudget genannt. Ein kleines ϵ bedeutet, dass der Privatsphäreschutz sehr hoch ist. Für diese Fall muss aber ein sehr starker privater Mechanismus angewandt werden, so dass die Nutzbarkeit der Daten sinken kann. Im Umkehrschluss steigt die Nutzbarkeit der Daten bei großem ϵ und die Daten werden weniger stark mit Rauschen belegt. Neben der Wahl des Privaten Mechanismus, spielt noch die Vertrauenswürdigkeit gegenüber der Datenbank eine Rolle. Falls diese vertrauenswürdig ist, kann Differential Privacy in einer globalen Variante verwendet werden, bei der alle Daten vor der Weitergabe zentral verrauscht werden. In diesem Fall kennt die Datenbank die realen Werte. Falls die Datenbank nicht vertrauenswürdig erscheint oder im Szenario der Datenweitergabe an eine dritte nicht vertrauenswürdige Partei, kann die lokale Variante von Differential Privacy, Local Differential Privacy, verwendet werden. Hierbei werden die einzelnen Datenpunkte individuell bei den Datensendenden verrauscht und erst dann weitergegeben. In diesem Fall kennt nur das Individuum die realen Daten.

Neben der theoretischen Erforschung der Differential Privacy Eigenschaft stehen inzwischen auch immer mehr Werkzeuge zur Verfügung die Technologie in der Praxis zu verwenden. Das OpenDP[71] Konsortium, gebildet aus Teilen von Microsoft und der Havard University, stellt zum Beispiel mit SmartNoise eine Programmbibliothek zur Verfügung, mit der sich DP in verschiedenen Anwendungen verwenden lässt. Allerdings löst diese Verfügbarkeit der Technologie noch nicht die diversen offenen Fragen in der Anwendung. So ist umstritten, welches Privatsphäre Budget ϵ verwendet werden soll und die Antwort ist auch Anwendungsfall abhängig. Eine wichtige Frage hierbei ist, wie das Privatsphäreschutz V.S. Nutzbarkeit Problem bewertet werden soll. Ein hoher Privatsphäreschutz durch ein kleines ϵ kann zur Akzeptanz der Technologie beitragen, schwächt allerdings den Forschungsnutzen der Daten. Die Forschungsseite würde vermutlich eher eine bessere Nutzbarkeit der Daten bevorzugen. Darüber hinaus gilt es auch die Differential Privacy Eigenschaft für nichttechnische Nutzende verständlich zu gestalten. Die Angabe eines konkreten Epsilon-Wertes wird vermutlich als eher weniger hilfreich betrachtet, um den Schutz der eigenen Daten bei einer Weitergabe zu bewerten. Hier könnten sich traditioneller Metriken wie k-anonymität oder l-diversity als durchaus leichter erklärbare Variante eignen.

Die zuvor beschriebene problematische Wahl von Epsilon, dessen Auswirkung und die verschiedenen Schwächen der traditionellen Technologien können eine große Unsicherheit für den Einsatz in Real-World Anwendungen bedeuten. Hierfür wäre eine Standardisierung verschiedener Privatsphäre wahrender Technologien denkbar, die den Technologie Einsatz regelt und einen Leitfaden bietet. Beispiele wie der zuvor in Kapitel III erwähnte HIPAA zeigen, dass die standardisierte reine Anonymisierung zwar eine rechtliche Sicherheit bietet, aber keineswegs ein Allheilmittel gegen Re-Identifizierung ist.[72] Im Entwurf für das Forschungsdatenzentrum soll der Versuch der Re-Identifizierung zwar sanktioniert[73] werden,

[71] Siehe: `https://opendp.org` (Letzter Zugriff: 10.03.2022)
[72] Siehe z.B. [JE18])
[73] Siehe § 303 e Abs. 6 SGB V

technische Maßnahmen werden dennoch nicht vorausgesetzt eine solche zu verhindern. Deswegen wäre die Pflicht jegliche geteilten oder veröffentlichten Daten, auf ein Re-Identifizierungsrisiko zu untersuchen und gegebenenfalls dieses Risiko durch den Einsatz zuvor beschrieben Technologien zu mindern, auf jeden Fall ein wichtiger Bestandteil eines solchen Leitfadens.

5.2 Mögliche Umsetzung

Auf Basis des zuvor erläuterten rechtlichen Rahmens und des aktuellen Standes der Technik wird hier eine mögliche Umsetzung für ein datenschutzkonformes Forschungsdatenzentrum aufgezeigt. Der Ansatz verfolgt die Prinzipien von **Transparenz**, **Kontrolle**, **Verfügbarkeit** und **Privatsphäreschutz**.

Abbildung 1 zeigt eine mögliche Architektur für ein datenschutzkonformes Forschungszentrum als Erweiterung des im DVG vorgeschlagenen Forschungsdatenzentrum.

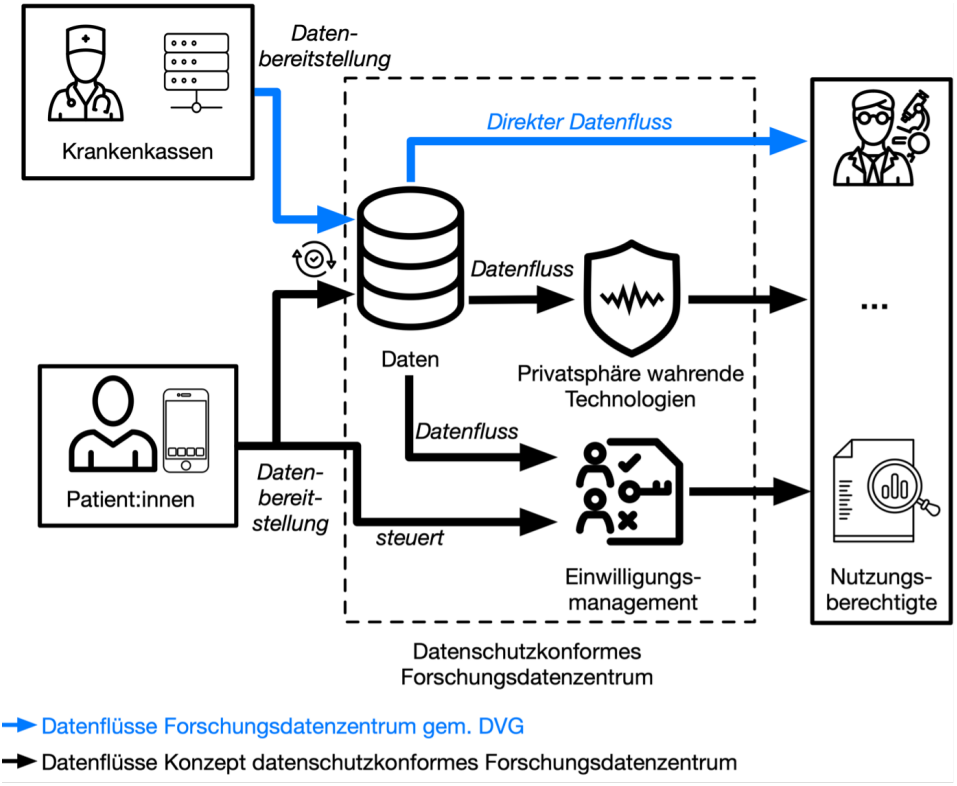

Abb. 1: Technische Architektur datenschutzkonformes Forschungszentrum

Der vorliegende Entwurf sieht im Wesentlichen vier Parteien vor. Zum einen Patient:innen, Krankenkassen, das datenschutzkonforme Forschungsdatenzentrum als Datentreuhänder und Nutzungsberechtigte. Die Daten im Forschungsdatenzentrum werden zum einen von den Datenlieferanten wie Krankenkassen, Ärzten oder Kliniken bereitgestellt, aber können auch beispielsweise direkt vom Patient:innen stammen (wie Lifestylefitnessdaten von Smartwatches). Denkbar wäre auch eine Anbindung an die Telematikinfrastruktur, die für die Elektronische Patientenakte (ePA) genutzt wird. Um Patient:innen Teilhabe zu ermöglichen, sollten Betroffene stets eine Übersicht haben, welche Daten im Forschungszentrum zur Verfügung stehen. Dies ermöglicht ein **transparentes** Verfahren für die Betroffenen. Für die Umsetzung bietet sich eine Smartphone Anwendung an, die auch weitere Verwaltungsmöglichkeiten für medizinische Daten bietet. Über eine solche App könnten Nutzer:innen auch **Kontrolle** über ihre Daten ausüben. So könnte ein Einwilligungsmanagement dazu dienen Daten explizit für Forschungsvorhaben freizugeben. Auf Basis dieser Einwilligungen wäre ein automatisierter Zugriff für die Forschungsinteressierten möglich. Wobei durch diesen Zugriff nur Daten, für die die explizite Einwilligungen von Betroffenen vorliegen, freigegeben werden. Diese Möglichkeit steigert die **Verfügbarkeit** der Daten. Für Massendaten, die beispielsweise in anonymisierter Form aufgrund der verschiedenen Rechtsgrundlagen genutzt werden können, sollten zusätzlich Privatsphäre wahrende Technologien verwendet werden, um auch das Risiko von Re-Identifizierungen zu senken. Somit wären die Forschungsinteressen gewahrt, bei einem gleichzeitig quantifizierbaren **Privatsphäreschutz** für Betroffene.

6 Fazit und Ausblick

Forschung auf bereits vorhandenen Versorgungsdaten in Deutschland hat ein großes Potential für die Verbesserung der medizinischen Behandlung. Im europäischen und internationalen Vergleich liegen die Möglichkeiten in Deutschland weit zurück. Mit dem Digitale-Versorgung-Gesetz wurde ein erster Schritt gegangen, um diese Forschung zu ermöglichen. Leider wurden dabei viele wichtige Fragen des Datenschutzes nicht geklärt. Die aktuell geplante Form der intransparenten Datenweitergaben in der Patient:innen keinerlei Mitspracherecht über die Nutzung ihrer Daten haben, erscheint gleich aus mehreren Gründen ungeeignet. Grundsätzlich stellt sich die Frage, ob dieser Teil des Gesetzes vor dem Bundesverfassungsgericht Bestand haben wird und ob es nicht europarechtswidrig ist. Weiter sehen die Autor:innen die Gefahr, möglicherweise langfristig die Akzeptanz und das Vertrauen der Bürger:innen zu verlieren. Im Vergleich zu dem im Gesetz skizzierten Ansatz, kann die im Paper vorgestellte Architektur die Transparenz und Mitsprache bei Patient:innen stärken. Dies hätte den weiteren Vorteil, dass sich über die Infrastruktur auch weitere Mechanismen, wie etwa eine Datenspende abbilden lassen.

Die aufgezeigten Problemfelder des Forschungsdatenzentrums erfassen nur einen Bruchteil der Fragen, die dessen Etablierung mit sich bringt. Ungeachtet der Transparenz- und Kontrollmöglichkeiten fehlen im aktuellen Stand des Digitale-Versorgung-Gesetzes verbindliche

Vorgaben zum Umgang mit den und damit einem fortgesetzten Schutz der bereitgestellten Daten. Während das japanische und amerikanische Recht sehr konkrete Vorgaben zum Schutz der betroffenen Personen gibt, fehlen diese in der DSGVO und dem DVG. Die Gefahr einer Re-Identifikation zunächst anonymer Daten ist zu befürchten, aufgrund des technischen Fortschritts in der Auswertung von Daten sowie durch die Verknüpfung unterschiedlicher Datensätze. Bisher bietet das geltende Recht allerdings keinerlei Lösung für dieses Defizit. Es verdeutlicht aber, dass technische Entwicklungen für die Gewährleistung des Datenschutzrechts immer wieder überprüft und neubewertet werden müssen.[74]

Mit Differential Privacy existiert eine Technologie, die nicht nur zur De-Identifikation von Daten geeignet ist, sondern gleichzeitig ein Maß für das Risiko einer Re-Identifizierung bietet. Dabei muss jedoch gesagt sein, dass es bisher zu wenig Erfahrung mit der Technologie gibt, um bereits Vorgaben in Form von konkreten Schwellwerten für Differential Privacy zu machen.

Danksagung

Diese Arbeit entstand in enger Zusammenarbeit mit dem KASTEL Projekt.

Literaturverzeichnis

[BK22] Becker, Ulrich; Kingreen, Thorsten: SGB V - Gesetzliche Krankenversicherung. Beck C. H., München, 2022.

[BSgD20] Bretthauer, Sebastian; Spiecker gen. Döhmann, Indra: Das Digitale-Versorgung-Gesetz als Einfallstor für eine Neujustierung von einstweiligem Rechtsschutz vor dem BVerfG und der Eingriffsqualität bei Datenverwendungen. JuristenZeitung, 75(20):990–996, 2020.

[Bu19] Bundesbeauftragte für den Datenschutz und die Informationsfreiheit (BfDI): , Stellungnahme des Bundesbeauftragten für den Datenschutz und die Informationsfreiheit zum Entwurf für das Digitale Versorgung-Gesetz,. https://www.bfdi.bund.de/SharedDocs/Downloads/DE/DokumenteBfDI/Stellungnahmen/2019/StgN_Digitale_Versorgung_Gesetz.pdf?__blob=publicationFile&v=4, 2019. zuletzt aufgerufen am 08.07.2022.

[Bu21] Buchner, Benedikt; Haber, Anna Christine; Hahn, Horst Karl; Prasser, Fabian; Kusch, Harald; Sax, Ulrich; Schmidt, Carsten Oliver: Das Modell der Datentreuhand in der medizinischen Forschung. Datenschutz Datensicherheit - DuD, 45(12):806–810, Dezember 2021.

[CR21] Calliess, Christian; Ruffert, Matthias: EUV/AEUV - Das Verfassungsrecht der Europäischen Union mit Europäischer Grundrechtecharta. Beck C. H., München, 2021.

[De17] Deutscher, Ethikrat: Big Data und Gesundheit : Datensouveränität als informationelle Freiheitsgestaltung : Stellungnahme. Deutscher Ethikrat, Berlin, 2017.

[74] Siehe hierzu: Spiecker gen. Döhmann, in: [Sp20], 489.

[Dw06] Dwork, Cynthia: Differential Privacy. In: Automata, Languages and Programming, Lecture notes in computer science, S. 1–12. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[D0] Däubler, Wolfgang; Wedde, Peter; Weichert, Thilo; Sommer, Imke: EU-DSGVO und BDSG - Kompaktkommentar. EU-Datenschutz-Grundverordnung (EU-DSGVO) - Neues Bundesdatenschutzgesetz (BDSG) - Weitere datenschutzrechtliche Vorschriften. Bund-Verlag GmbH, Nördlingen, 2020.

[Gi21] Gierschmann, Sibylle: Gestaltungsmöglichkeiten durch systematisches und risikobasiertes Vorgehen – Was ist schon anonym? Zeitschrift für Datenschutz, 9:482–486, 2021.

[GLF18] Geminn, Christian; Laubach, Anne; Fujiwara, Shizuo: Schutz anonymisierter Daten im japanischen Datenschutzrecht. Zeitschrift für Datenschutz, S. 413–419, 2018.

[JE18] Janmey, V.; Elkin, P. L.: Re-Identification Risk in HIPAA De-Identified Datasets: The MVA Attack. AMIA Annu Symp Proc, 2018:1329–1337, 2018.

[KB20] Kühling, Jürgen; Buchner, Benedikt: Datenschutz-Grundverordnung/ Bundesdatenschutzgesetz. C.H.Beck, München, 2020.

[KS20] Kühling, Jürgen; Schildbach, Roman: Die Reform der Datentransparenzvorschriften im SGB V. Neue Zeitschrift für Sozialrecht, 2:41–50, 2020.

[K0] Kühling, Jürgen: Gesundheitsdatenschutzrecht im Zeitalter von „Big Data“. Datenschutz Datensicherheit - DuD, 44(3):182–188, März 2020.

[K1] Kühling, Jürgen: Der datenschutzrechtliche Rahmen für Datentreuhänder. Datenschutz Datensicherheit - DuD, 45(12):783–788, Dezember 2021.

[La16] Ladeur, Karl-Heinz: „Big Data“ im Gesundheitsrecht – Ende der „Datensparsamkeit“? Datenschutz Datensicherheit - DuD, 40(6):360–364, Juni 2016.

[Le13] Lejeune, Mathias: Datenschutz in den Vereinigten Staaten von Amerika. CR, S. 755–760, 2013.

[LLV07] Li, Ninghui; Li, Tiancheng; Venkatasubramanian, Suresh: t-Closeness: Privacy Beyond k-Anonymity and l-Diversity. In: 2007 IEEE 23rd International Conference on Data Engineering. S. 106–115, 2007.

[Ma06] Machanavajjhala, A.; Gehrke, J.; Kifer, D.; Venkitasubramaniam, M.: L-diversity: privacy beyond k-anonymity. In: 22nd International Conference on Data Engineering (ICDE'06). S. 24–24, 2006.

[Me21] Meyer, Stephan: Landesrechtliche Legaldefinitionen der „Anonymisierung“ im Anwendungsbereich der DS-GVO. Zeitschrift für Datenschutz, 12:669–673, 2021.

[MH20] Martini, Mario; Hohmann, Matthias: Der gläserne Patient: Dystopie oder Zukunftsrealität? Neue Juristische Wochenzeitschrift, 49:3573–3578, 2020.

[NAC07] Nergiz, Mehmet Ercan; Atzori, Maurizio; Clifton, Chris: Hiding the presence of individuals from shared databases. In: Proceedings of the 2007 ACM SIGMOD international conference on Management of data - SIGMOD '07. ACM Press, New York, New York, USA, 2007.

[PP21] Paal, Boris P.; Pauly, Daniel A.: Datenschutz-Grundverordnung Bundesdatenschutzgesetz. Beck C. H., München, 2021.

[RG21] Roßnagel, Alexander; Geminn, Christian L.: Vertrauen in Anonymisierung. Zeitschrift für Datenschutz, 9:487–490, 2021.

[Ro18] Rolfs, Christian; Giesen, Richard; Kreikebohm, Ralf; Udsching, Peter: BeckOK Sozialrecht. Verlag C.H. Beck, München, 2018.

[Ro19] Roßnagel, Alexander: Datenschutz in der Forschung. Zeitschrift für Datenschutz, S. 157–159, 2019.

[Ro21] Roßnagel, Alexander: Grundrechtsschutz in der Datenwirtschaft. Zeitschrift für Rechtspolitik, 6:173–175, 2021.

[SgDB20] Spiecker gen. Döhmann, Indra; Bretthauer, Sebastian: , Schutzlos in Karlsruhe: Neue Maßstäbe im einstweiligen Rechtsschutz und im Datenschutz vor dem Bundesverfassungsgericht. https://verfassungsblog.de/schutzlos-in-karlsruhe/, 2020. zuletzt aufgerufen am 08.07.2022.

[SHD19] Simitis, Spiros; Hornung, Gerrit; Döhmann, Indra Spiecker: Datenschutzrecht - DSGVO mit BDSG. Nomos Verlagsgesellschaft, Baden-Baden, 2019.

[Sp17] Spiecker gen. Döhmann, Indra: Big und Smart Data: Zweckbindung zwecklos? Spektrum der Wissenschaft, 2017(1):56–62, 2017.

[Sp20] Specht-Riemenschneider, Louisa; Buchner, Benedikt; Heinze, Christian; Thomsen, Oliver: Festschrift für Jürgen Taeger - IT-Recht in Wissenschaft und Praxis. Fachmedien Recht und Wirtschaft, 2020.

[St17] StiftungDatenschutz: Big Data und E-Health. Erich Schmidt Verlag GmbH Company, Berlin, 2017.

[Sw00] Sweeney, Latanya: Simple demographics often identify people uniquely. Health (San Francisco), 671:1–34, 2000.

[Sw02] Sweeney, Latanya: K-ANONYMITY: A MODEL FOR PROTECTING PRIVACY. Internat. J. Uncertain. Fuzziness Knowledge-Based Systems, 10(05):557–570, Oktober 2002.

[SW17] Schantz, Peter; Wolff, Heinrich Amadeus: Das neue Datenschutzrecht. C.H.Beck, München, 2017.

[We20] Weichert, Thilo: "Datentransparenz" und Datenschutz. Medizinrecht, 38(7):539–546, Juli 2020.

[Zu21] Zuo, Z.; Watson, M.; Budgen, D.; Hall, R.; Kennelly, C.; Al Moubayed, N.: Data Anonymization for Pervasive Health Care: Systematic Literature Mapping Study. JMIR Med Inform, 9(10):e29871, Oct 2021.