Benedikt Buchner, Anna C. Haber, Horst K. Hahn, Harald Kusch, Fabian Prasser, Ulrich Sax, Carsten O. Schmidt*

Das Modell der Datentreuhand in der medizinischen Forschung

Der Konflikt zwischen Datenschutz und Forschungsfreiheit ist so alt wie das Datenschutzrecht selbst. In vielen Fällen wird Datenschutz als ein Hemmschuh für die Forschung wahrgenommen, gerade auch im Bereich der medizinischen Forschung. Ein aktuell viel diskutierter Lösungsansatz, diesen Konflikt zu entschärfen, ist das Modell der Datentreuhand. Ob und wie ein solches Modell nutzbar gemacht werden kann, um für die medizinische Forschung einen leichteren und großzügigeren Zugang zu personenbezogenen Daten zu eröffnen, soll im Folgenden erörtert werden.

1 Forschungsfreiheit versus Datenschutz?

In einem Editorial dieser Zeitschrift wurde das Spannungsfeld Datenschutz und Forschungsfreiheit schon 1999 zu den „Evergreens der Datenschutzdiskussion" gezählt.[1] Die Diskussion, ob eine zu strenge Datenschutzregulierung die Forschung beschränkt oder gar verhindert, reicht zurück bis in die 70er Jahre, als das Bundes- und die Landesdatenschutzgesetze nach und nach in Kraft traten.[2] Für die medizinische Forschung wies der Wissenschaftsrat bereits in einer Stellungnahme aus dem Jahr 1982 darauf hin, dass sich die datenschutzrechtlichen Probleme hier „mit besonderer Schärfe" stellen würden. Naturgemäß sei es in der Medizin besonders schwierig, sich auf den Erlaubnistatbestand der Einwilligung zu stützen, da diese streng zweckgebunden zu erfolgen hat, der Zweck bei medizinischen Forschungsvorhaben jedoch nicht immer von vornherein klar umrissen werden kann.[3] Die Rede ist gar von einem „Prozess resignierender Selbstbeschränkung" der Wissenschaftler, weil die Forschung zunehmend darauf verzichte, bestimmte Fragen überhaupt noch in Angriff zu nehmen, um nicht mit Datenschutzregelungen in Konflikt zu geraten.[4]

In der Tat hatte die Datenschutzgesetzgebung in ihren Anfangszeiten die Belange der Forschung kaum im Blick. Spezifisch forschungsbezogene Regelungen, die auf einen Ausgleich zwischen Datenschutz und Forschungsfreiheit abzielten, fanden sich zunächst nur in den Landesdatenschutzgesetzen von Baden-Württemberg, Hessen, Nordrhein-Westfalen und Rheinland-Pfalz.[5] Insgesamt war das Spannungsverhältnis zwischen Datenschutz und Forschung ein Thema, das die Datenschutzgesetzgebung in ihren Anfangszeiten noch nicht hinreichend berücksichtigt hatte.

Spätestens seit den 90er Jahren trägt dieser Befund einer weitestgehenden Forschungsblindheit des Datenschutzrechts allerdings nicht mehr. Der mit der Novellierung des BDSG 1990 ausgelöste „legislative Schub forschungsfreundlicher Regelungen"[6] hat sich bis heute fortgesetzt – nunmehr in erster Linie auf europäischer Ebene. Die Datenschutz-Grundverordnung enthält eine Vielzahl von Regelungsansätzen, die allesamt darauf abzielen, eine Datenverarbeitung zu wissenschaftlichen Forschungszwecken zu ermöglichen. Dies gilt zuallererst für die Lockerung des Zweckbindungsgrundsatzes: Art. 5 Abs. 1 lit. b Halbsatz 2 DSGVO stellt insoweit die Fiktion auf, dass eine Weiterverarbeitung von Daten zu wissenschaftlichen Forschungszwecken mit dem ursprünglich verfolgten Zweck einer Datenverarbeitung nicht unvereinbar ist und räumt damit eine ganz wesentliche datenschutzrechtliche Hürde zugunsten der Forschung beiseite. Von zentraler Bedeutung gerade für die medizinische Forschung ist auch die Öffnungsklausel des Art. 9 Abs. 2 lit. j DSGVO, die es den Mitgliedstaaten erlaubt, Ausnahmen vom grundsätzlichen Verbot einer Verarbeitung von besonders schutzwürdigen Daten wie Gesundheitsdaten zu normieren, wenn eine Datenverarbeitung für wissenschaftliche Forschungszwecke erforderlich ist. Auch der Erlaubnistatbestand der Einwilligung erfährt für den Forschungsbereich dadurch eine großzügigere Handhabe, dass die DSGVO in ihrem Erwägungsgrund 33 die Rechtsfigur des sog. *broad consent* für die wissenschaftliche Forschung an-

* Prof. Dr. Benedikt Buchner, Institut für Informations-, Gesundheits- und Medizinrecht, Universität Bremen; Dipl.-Math. Anna Christine Haber, AG Medizininformatik, Berlin Institute of Health at Charité – Universitätsmedizin Berlin; Prof. Dr.-Ing. Horst Karl Hahn, Fraunhofer-Institut für Digitale Medizin MEVIS; Prof. Dr. Fabian Prasser, AG Medizininformatik, Berlin Institute of Health at Charité – Universitätsmedizin Berlin; Dr. Harald Kusch, Institut für Medizinische Informatik, Universitätsmedizin Göttingen; Prof. Dr. Ulrich Sax, Institut für Medizinische Informatik, Universitätsmedizin Göttingen; Prof. Dr. Carsten Oliver Schmidt, Institut für Community Medicine, Universitätsmedizin Greifswald.

Diese Arbeit ist im Rahmen des NFDI4Health-Konsortiums entstanden (www.nfdi4health.de). Wir danken der Deutschen Forschungsgemeinschaft (DFG) für die finanzielle Unterstützung – NFDI 13/1.

1 *Bizer*, Bewährte Rezepte – Editorial, DuD 1999, 374.

2 Siehe etwa schon *Ziegler-Jung*, Datenschutz und Datenzugang der Forschung im Gesundheitsbereich, in: *Brennecke/Greiser/Paul/Schach*, Datenquellen für Sozialmedizin und Epidemiologie (1981), S. 37.

3 *Wissenschaftsrat*, Stellungnahme zu Forschung und Datenschutz vom 5.11.1982, Drs. 5900/82, S. 18 f.

4 *Wissenschaftsrat* (Fn. 3), S. 8 f.

5 Vgl. *Wissenschaftsrat* (Fn. 3), S. 26; *Ziegler-Jung* (Fn. 2), S. 40 f.

6 *Bizer*, DuD 1999, 374.

erkennt: Als Ausnahme vom Grundsatz der Zweckbestimmtheit einer Einwilligung soll für bestimmte Bereiche wissenschaftlicher Forschung auch eine deutlich allgemeiner gehaltene Einwilligung zulässig sein. Die genaue Zielsetzung von neuen Forschungsprojekten muss also nicht von vornherein genau festgelegt sein, um sich auf den Erlaubnistatbestand der Einwilligung für eine Verarbeitung personenbezogener Daten stützen zu können. Im Rahmen der Medizininformatik-Initiative wurden die Optionen für den Broad Consent zur Nutzung von Patientendaten aus Krankenhäusern ausgelotet und mit dem Arbeitskreis Medizinischer Ethik-Kommissionen abgestimmt.[7] Die Implementierung findet derzeit in den deutschen Universitätsklinika statt.

In Umsetzung des Regelungsspielraums, den die DSGVO dem nationalen Gesetzgeber für den Forschungsbereich lässt, findet sich auch im nationalen Recht eine Vielzahl von Privilegierungstatbeständen für die wissenschaftliche Forschung. § 27 Abs. 1 Satz 1 BGB erlaubt in Form einer Interessenabwägungsklausel eine Datenverarbeitung für Forschungszwecke, „wenn die Verarbeitung zu diesen Zwecken erforderlich ist und Interessen des Verantwortlichen an der Verarbeitung die Interessen der betroffenen Person an einem Ausschluss der Verarbeitung erheblich überwiegen". Daneben finden sich zahlreiche bereichsspezifische Forschungsklauseln, etwa für Vorhaben der wissenschaftlichen Forschung im Sozialleistungsbereich (§ 75 SGB X) oder auch Forschungsklauseln im Landeskrankenhausrecht, die eine Verarbeitung von Patientendaten für wissenschaftliche medizinische Forschungsvorhaben unter bestimmten Voraussetzungen erlauben.[8]

Mit Blick auf diese und viele weitere forschungsbezogene Regelungsansätze des geltenden Datenschutzrechts lässt sich also festhalten: Datenschutz und Forschung mögen stets in einem gewissen Spannungsverhältnis stehen, weil sie in Teilen gegenläufige Zielsetzungen verfolgen. Das Datenschutz*recht* trägt diesem Umstand jedoch seit langem Rechnung und sieht eine Vielzahl von Regelungsinstrumentarien vor, um die Zielsetzungen von Datenschutz und Forschung weitestmöglich in Einklang zu bringen. Entsprechend muss es in der aktuellen Diskussion zuallererst darum gehen, wie auf der Basis des geltenden, durchaus forschungsfreundlich angelegten Datenschutzrechts Lösungsansätze entwickelt werden können, die den Forschungsbedürfnissen unter gleichzeitiger Wahrung informationeller Selbstbestimmung angemessen Rechnung tragen. Dies wäre dann auch ganz im Sinne des Wissenschaftsrats, der schon 1982 für die Vereinbarkeit von Datenschutz und Forschungsfreiheit eine Unterscheidung eingefordert hat „zwischen Problemen, die nur der Gesetzgeber lösen kann, und Problemen, die durch eine vernünftige Auslegung geltenden Rechts bewältigt werden können." Eben *letztere* Probleme gilt es zunächst einmal zu lösen – bevor zum wiederholten Male nach dem Gesetzgeber gerufen wird. Dies gilt auch für die Frage, ob und wie das Modell der Datentreuhand als möglicher Lösungsansatz im medizinischen Forschungsbereich berücksichtigt werden kann und soll.

2 Das Modell der Datentreuhand

Das Modell der Datentreuhand hat in den letzten Jahren einen regelrechten Hype erfahren, nicht nur in der rechtswissenschaftlichen Literatur und in zahlreichen Datenschutzkonzepten der Verbundforschung, sondern auch bei vielen Gremien und politischen Entscheidungsträgern. So sieht etwa die Datenethikkommission in ihrem Gutachten von 2019 in Datentreuhandsystemen ein „großes Potential" und empfiehlt, Forschung und Entwicklung in diesem Bereich intensiv zu fördern.[9] Auch dem Rat für Informationsinfrastrukturen erscheint die Idee eines Systems von Datentreuhandstellen „überaus sinnvoll für die Entwicklung einer Datenstrategie", soweit diesem ein „Grundkonzept der gleichermaßen effektiven wie rechtskonformen und im Sinne der Datensouveränität nachvollziehbaren Nutzung von qualitätsgesicherten Datenbeständen" zugrunde liegt.[10] Schließlich sprach auch die letzte Bundesregierung in ihrer Datenstrategie 2021 Datentreuhändern eine wichtige Funktion zu, Datenzugang und -austausch zu sichern und dabei gleichzeitig auch eine Einhaltung des geltenden Datenschutzrechts zu gewährleisten.[11] Ein solcher Datentreuhänder im besten Sinne ist auch das durch § 303d-e SGB V sowie durch die Neufassung der Datentransparenzverordnung (DaTraV) definierte Forschungsdatenzentrum (FDZ).

2.1 Die Idee der Datentreuhand

Das Modell der Datentreuhand mag aktuell in aller Munde sein, neu ist diese Idee allerdings keineswegs. So ging auch schon der Wissenschaftsrat in der eingangs erwähnten Stellungnahme aus 1982[12] auf die Überlegung ein, ob nicht Datentreuhänder eingesetzt werden könnten, um einerseits mittels Anonymisierung von Daten diese gegenüber einem unbefugten Zugriff zu sichern, andererseits aber gleichwohl die Möglichkeit zu erhalten, für wissenschaftliche Bedürfnisse den Personenbezug von Daten gegebenenfalls wiederherzustellen.[13] Und auch in dieser Zeitschrift ist schon vor über 30 Jahren ein konkreter Vorschlag präsentiert worden, wie die Rolle von Datentreuhändern genutzt werden kann, um eine Datennutzung zu wissenschaftlichen Forschungszwecken zu ermöglichen. Bereits damals ging die Grundidee dahin, dass Treuhänder personenbezogene Daten eines Betroffenen, die an verschiedenen Stellen anfallen, zusammenführen, analysieren und dann an Forscher anonym übermitteln.[14] Ebenso wurden in verschiedenen Rechtsgutachten konkrete Fragestellungen der Datentreuhand erörtert und Spielräume ausgelotet, etwa das Zusammenspiel mit dem Zeugnisverweigerungsrecht von Ärzten und dem Beschlagnahmeschutz für gewisse Daten und Berufsgruppen.[15] So wurde bereits mit Einführung der elektroni-

7 *Bild/Bialke/Buckow/Ganslandt/Ihrig/Jahns/Merzweiler/Roschka/Schreiweis/Stäubert/Zenker/Prasser*, 2020. Towards a comprehensive and interoperable representation of consent-based data usage permissions in the German medical informatics initiative. BMC Med Inform Decis Mak 20, 103. https://doi.org/10.1186/s12911-020-01138-6.

8 Für einen Überblick über die Forschungsklauseln in den einzelnen Landeskrankenhausgesetzen s. *Pollmann* in *Buchner* (Hrsg.), Datenschutz im Gesundheitwesen (27. EL 2021), C/5.2.

9 *Datenethikkommission*, Gutachten (2019), S. 21 und 133 ff.

10 *Rat für Infrastrukturen*, RfII-Stellungnahme „Datentreuhandstellen gestalten – zu Erfahrungen der Wissenschaft" (2020), S. 5.

11 *Bundesregierung*, Datenstrategie (2021), S. 34.

12 *Wissenschaftsrat* (Fn. 3), S. 31 f.

13 Der Begriff der Anonymisierung und seine Limitierung und inhaltliche Ausgestaltung sind keinesfalls „gesetzt" (wenn er auch streng mathematisch definiert werden kann), sondern unterliegen einem steten Wandel und werden kritisch und kontrovers diskutiert. Für eine kurze Einordnung der Anonymisierung und insbesondere der Anonymisierbarkeit von Datenbeständen in die Thematik der Datentreuhand s. u. 4 (im Sinne eines Ausblicks auf das Leistbare).

14 *Erfa-Kreis*, DuD 1989, 125, 129 (zitiert nach *Bizer*, DuD 1999, 392, 395).

15 S. etwa *Dierks*, Rechtsgutachten zur elektronischen Datentreuhänderschaft (2008); abrufbar unter www.tmf-ev.de/Themen/Projekte/V052_01_Daten-

schen Gesundheitskarte der Kreis derjenigen, in deren Händen sich sensible Gesundheitsdaten befinden und für die z. B. der Beschlagnahmeschutz zu gelten habe, erweitert – in diesem Fall auf die Patienten selbst. Gleiches wird im Zusammenhang mit der Datentreuhänderschaft nun seit Längerem für sog. Dienstleister (konkret hier den Treuhänder) sowie eventuelle nachrangige Dienstleister diskutiert. Die Thematik nimmt noch einmal Fahrt auf im Zuge der Etablierung der (einrichtungsübergreifenden) elektronischen Patientenakte (ePA), die ab dem 01.01.2023 eine Erweiterung um die Forschungsnutzung erhält.[16]

2.2 Datentreuhandmodelle in der medizinischen Forschung

Die Beispiele für Treuhandmodelle, wie sie im Bereich der medizinischen Forschung heute schon praktiziert werden, sind zahlreich. Im Folgenden sollen nur ein paar dieser Modelle exemplarisch vorgestellt werden.

Generisches Datenschutzkonzept der TMF: Im Rahmen der Förderung der Kompetenznetze in der Medizin[17] tauchte bereits Ende der 1990er die Frage nach der Vereinbarkeit von Forschung und Datenschutz systematisch und mit besonderem Nachdruck in mehreren medizinischen Disziplinen auf. Als Reaktion darauf wurden von übergreifenden methodischen Arbeitsgruppen unter dem Dach der TMF generische Datenschutzkonzepte entwickelt[18], mit dem AK Wissenschaft und der AG Technik der Landesdatenschutzbeauftragten abgestimmt und in den folgenden Jahren in vielen Verbundforschungsvorhaben umgesetzt[19]. Diese Konzepte stoßen angesichts der DSGVO an ihre Grenzen und werden seitens der TMF derzeit grundlegend überarbeitet.

Das generische Datenschutzkonzept der TMF fußt vereinfacht gesagt auf der Reduzierung des Reidentifizierungsrisikos durch Aufteilung des Datenbestandes in identifizierende Daten (IDAT) und medizinische Daten (MDAT). Die Zuordnung der Daten erfolgt über eine so genannte Patientenliste, die an einen Pseudonymisierungsdienst gekoppelt ist. So enthalten die MDAT ursprünglich jeweils noch eine Patientennummer oder den Namen, die Patientenliste ergänzt ein synonymarmes und homonymes Pseudonym (PSN). Die MDAT-Tabelle wiederum enthält nur noch das PSN als Primärschlüssel. Nur über die Patientenliste ist eine anlassbezogene De-Pseudoymisierung möglich. Daher ist die Patientenliste an einer vertrauenswürdigen Stelle zu betreiben. Auch hier wurde der Begriff des „Treuhänders" verwendet. Nach anfänglichen Überlegungen, ob nicht die TMF diesen Treuhänderdienst betreiben solle, wurde dieser Auftrag u. a. wegen der Sicherheitsvermutung eines besseren Beschlagnahmeschutzes an Betreiber von IT-Komponenten in klinischen Einrichtungen bzw. andere Rechenzentren vergeben. Neben dem reinen Serverhosting sind mit der Treuhänderschaft auch Aufgaben gekoppelt wie die Löschung bzw. Markierung fehlerhafter Einträge und die Zusammenführung synonymer Datensätze. Weiterhin kann hier auch das Vorliegen einer Einwilligung bzw. der Widerruf einer Einwilligung dokumentiert werden.

MOSAIC-Projekt: Dem gerade skizzierten Modell folgend werden im Rahmen des MOSAIC-Projekts Daten im kardiologischen Kontext (Patient:innen mit Transkatheter-Aortenklappenimplantation, DOI: 10.5334/ojb.65) in der Universitätsmedizin Göttingen unter Einbeziehung einer unabhängigen Treuhandstelle in Greifswald[20] prozessiert.[21] Hierbei können klinisch-phänotypische Daten, Bioproben-Daten und personenbezogene Daten zunächst datenschutzkonform getrennt erhoben und verwaltet werden. Für integrierte wissenschaftliche Analysen (z.B. DOI: 10.1093/eurheartj/ehaa033) können diese Daten ebenfalls unter Berücksichtigung der gesetzlichen Rahmenbedingungen wieder zusammengeführt und bearbeitet werden. Aus diesem komplexen Setup ergeben sich neue Anforderungen an die Bearbeitung der Datenintegrationsschritte, die u. U. durch spezielle Schulungen der Datenbereitsteller und -nutzer unterstützt werden müssen. Im beschriebenen Anwendungsfall wird die Datenintegration (in Form von „Extract, Transform, Load" (ETL) –Prozessen) einschließlich der „Application programming Interface" (API)-basierten Abfragen an die Treuhandstelle weitgehend automatisiert, so dass zum Zeitpunkt der Etablierung zwar ein höherer Entwicklungsaufwand entsteht, in der weiteren oft langjährigen Verwendung der Daten diese Prozesse aber effizient nachnutzbar sind.

Epidemiologische Kohortenstudien: Auch wenn formell nicht der Begriff der Treuhandstelle genutzt wird, sind analoge organisatorische Einteilungen in großen epidemiologischen Kohortenstudien wie der *Study of Health in Pomerania (SHIP)* bereits lange umgesetzt. Dies betrifft z. B. seit Studienbeginn 1997 organisatorisch komplett abgetrennte Einheiten zur Verwaltung von personenidentifizierenden Daten (IDAT) im Probandenmanagement (PBM) und von medizinischen Daten (MDAT) im Datenmanagement (DM). Das PBM arbeitet auf einem vom Internet komplett abgeschirmten Intranet, um externen Angreifern keine Angriffsfläche auf die IDAT zu bieten. Es bestehen auch intern keine direkten Schnittstellen zum DM. Das PBM koordiniert neben der Verwaltung der IDAT und der Probandenrekrutierung auch die Kopplung externer Datenbestände, z. B. von Routinedaten aus der Gesundheitsversorgung. Die Aufgabe des PBM beschränkt sich dabei auf die korrekte Zuordnung von IDAT unter Beachtung der erteilten Einverständnisse. MDAT aus Datenkopplungen werden hingegen direkt an das DM übergeben. Anregungen aus anderen Konzepten, wie dem oben dargelegten generischen TMF-Datenschutzkonzept, führten zur weiteren Verbesserung des Datenschutzes, etwa durch die differenzierte Verwendung von im PBM generierten Pseudonymen. So werden für verschiedene Erhebungsbereiche der SHIP eigene Pseudonyme verwendet, aber auch für externe Datenübergaben von MDAT projektspezifische Pseudonyme generiert, um das Risiko unerwünschter Datenkopplungen zu senken. Das PBM nimmt also faktisch treuhänderische Aufgaben wahr im Sinne eines Vermittlers datenschutzrechtlicher Einwilligungen. Ein wichtiger Bestandteil davon ist die Übergabe der aktuellen Stände der Einverständnis-

treuhaenderdienst_l.aspx.

16 S. insb. § 363 SGB V, der die freiwillige Weitergabe von Daten der ePA zu bestimmten Forschungszwecken an das FDZ regelt.

17 *Uhlig*, 15 Jahre Kompetenznetze in der Medizin, Editorial Bundesgesundheitsblatt 4/2016.

18 *Pommerening/Sax/Müller/Speer/Ganslandt/Drepper et al.*, Integrating eHealth and Medical Research: The TMF Data Protection Scheme. In: *Blobel/Pharow/Zvarova/Lopez* (ed.). eHealth: Combining Health Telematics, Telemedicine, Biomedical Engineering and Bioinformatics to the Edge (2008), S. 5 ff.

19 *S. Helbing et al.*, A data protection scheme for medical research networks. Review after five years of operation, Methods Inf Med 2010, DOI: 10.3414/ME09-02-0058.

20 https://www.ths-greifswald.de/.

21 Angelehnt an die Umsetzung im Deutschen Zentrum für Herz-Kreislauf-Forschung e.V., DZHK, https://www.ths-greifswald.de/projekte/dzhk/.

se an das DM, um auf dieser Basis die Weichen für eine einwilligungskonforme Nutzung der Daten und ggfs. deren Löschung sicherzustellen. Der Prozess von der Datenbeantragung bis zur Übergabe der MDAT wird durch eine Transferstelle koordiniert. Der Einklang beabsichtigter Datennutzungen mit den Zielen der SHIP wird dabei in einem formellen Begutachtungsprozess durch den Forschungsverbund Community Medicine im Rahmen monatlicher Vorstandssitzungen geprüft.

Forschungsdatenzentrum (FDZ) des Bundesgesundheitsministeriums: Ziel des Aufbaus des FDZ ist es, einen geschützten und vertrauenswürdigen Datenraum für „die Nutzung der Abrechnungsdaten der gesetzlich Krankenversicherten zur Präventions- und Versorgungsforschung und zur Steuerung des Gesundheitswesens“[22] zu schaffen, der dann einem Kreis berechtigter Institutionen zugänglich gemacht wird. Unter anderem übernimmt das FDZ als Datentreuhänder (wenngleich dieser Begriff in den entsprechenden Gesetzen §303a-f SGB V und der DaTraV nicht verwendet wird) die Aufgaben der Datenaufbereitung, Daten-Qualitätskontrolle, Überprüfung der Nutzungsanträge und Verfügbarmachung. Besonders hervorzuheben ist zudem die Aufgabe, jeweils das „spezifische Reidentifikationsrisiko in Bezug auf die durch Nutzungsberechtigte nach § 303e [SGB V] beantragten Daten zu bewerten und unter angemessener Wahrung des angestrebten wissenschaftlichen Nutzens durch geeignete Maßnahmen zu minimieren“ (§ 303d Abs. 1 Ziff. 5 SGB V).

Wie bereits erwähnt, ist zudem ab 2023 entsprechend § 363 SGB V vorgesehen, dass Versicherte auf Basis einer informierten Einwilligung individuell Daten aus ihrer ePA ganz oder teilweise für bestimmte Forschungszwecke über das FDZ verfügbar machen können. Dies gilt bis zum Widerruf, welcher zur Löschung der Daten im FDZ führt. In jedem Falle dürfen jedoch die bereits „übermittelten und für konkrete Forschungsvorhaben … verwendeten Daten … weiterhin für diese Forschungsvorhaben verarbeitet werden“ (§ 363 Abs. 6 SGB V).

3 Rechtliche Rahmenbedingungen

Geht es um die Möglichkeit einer Datenverarbeitung zu Forschungszwecken, ist ganz offensichtlich die Wahrnehmung vorherrschend, dass es hierfür im Regelfall einer Einwilligung der betroffenen Personen bedarf, die jedoch nicht immer – oder zumindest nicht immer rechtssicher – als Legitimationsgrundlage zur Verfügung steht. Eine Lösung über das Datentreuhandmodell wird vor diesem Hintergrund vor allem in zwei Varianten angedacht:

Zum einen wird vorgeschlagen, dass ein Datentreuhänder als Vermittler einer datenschutzrechtlichen Einwilligung eingesetzt wird. Der Datentreuhänder agiert hier als Vertrauensperson von Betroffenen und soll diesen dabei helfen, ihr informationelles Selbstbestimmungsrecht (in seiner Ausprägung als Einwilligung) möglichst effektiv auszuüben. Der Datentreuhänder handelt hier also im Auftrag der betroffenen Personen und nimmt deren datenschutzrechtliche Interessen gegenüber Dritten wahr, indem er stellvertretend für die betroffene Person eine Einwilligung erteilt, beschränkt oder auch widerruft.[23]

Zum anderen findet sich insbesondere für den Bereich der medizinischen Forschung aber auch der Lösungsvorschlag, ein Datentreuhandmodell *als Alternative* zum Erlaubnistatbestand der Einwilligung einzuführen. Etabliert werden soll eine gesetzliche Regelung von Datentreuhändern für medizinische Daten, die einen Erlaubnistatbestand für die Datenverarbeitung zu Zwecken medizinischer Forschung schafft.[24]

3.1 Einwilligung als primäre Legitimationsgrundlage?

Bei beiden Lösungsansätzen stellt sich allerdings für den Bereich der Forschungsdatenverarbeitung vorgeschaltet die Frage, ob deren Prämisse, nämlich die zentrale Rolle der Einwilligung als Erlaubnistatbestand, so überhaupt zutreffend ist. In der Praxis ist es sicherlich bislang so, dass sich die medizinische Forschung zuallererst auf die Einwilligung als Legitimationsgrundlage für eine Nutzung personenbezogener Daten stützt bzw. stützen möchte. Die Frage ist allerdings, ob diese Praxis so zwingend ist oder ob sie möglicherweise die Spielräume, die das Datenschutzrecht für die Forschungsdatenverarbeitung vorsieht, noch gar nicht ausschöpft.

Wirft man einen Blick auf die zentralen Privilegierungstatbestände des Datenschutzrechts für eine Forschungsdatenverarbeitung, so lässt sich aus diesen jedenfalls nicht herauslesen, dass es gerade die Einwilligung der betroffenen Personen sein soll, die eine Datenverarbeitung zu Forschungszwecken zuallererst legitimiert. Die Öffnungsklausel des Art. 9 Abs. 2 lit. j DSGVO spricht die Einwilligung als mögliche oder gar primäre Legitimationsgrundlage überhaupt nicht an und § 27 Abs. 1 BDSG, die zentrale allgemeine Forschungsklausel im nationalen Recht, stellt explizit klar, dass selbst eine Verarbeitung von besonders sensiblen Daten wie Gesundheitsdaten für wissenschaftliche Forschungszwecke auch ohne Einwilligung zulässig ist, wenn eine Interessenabwägung zugunsten der Forschungsinteressen ausfällt.

3.2 Datentreuhand als Schutzmaßnahme

Was Art. 9 DSGVO und § 27 BDSG allerdings vorsehen, ist eine Absicherung der Grundrechte und Interessen der betroffenen Personen durch *angemessene und spezifische Schutzmaßnahmen*. Überhaupt kann eine Interessenabwägung nur dann zugunsten der medizinischen Forschung ausfallen, wenn die jeweiligen Forscher oder Forschungsinstitutionen alle zur Verfügung stehenden Schutzmaßnahmen ergriffen haben, um das informationelle Selbstbestimmungsrecht der betroffenen Personen so weit wie möglich abzusichern. Und aus eben diesem Grund kann dann auch der Datentreuhand eine zentrale Funktion als „angemessene und spezifische Maßnahme“ im Sinne der datenschutzrechtlichen Vorgaben zukommen. Mittels eines Datentreuhänders kann gewährleistet werden, dass personenbezogene Daten für die Forschung selbst nur in anonymisierter Form zur Verfügung gestellt werden (und damit insoweit datenschutzrechtlich irrelevant sind), gleichzeitig aber die Möglichkeit erhalten bleibt, Daten aus unterschiedlichen Quellen und/oder unterschiedlichen Zeiträumen zu ein und derselben Person dank des Zuordnungswis-

22 *Bundesregierung* (Fn. 11), S. 30 f.

23 S. grundsätzlich zu diesem Modell den Beitrag von *Kühling* in diesem Heft.

24 Siehe zu diesem Ansatz *Bankertz/Riemenschneider* in boell.brief Juli 2021, Neue Modelle ermöglichen – Regulierung für Datentreuhänder, S. 10; www.boell.de/de/2021/07/09/neue-modelle-ermoeglichen.

sens beim Treuhänder stets korrekt zu verknüpfen.[25] Das Modell der Datentreuhand trägt damit sowohl der datenschutzrechtlichen Maxime einer Erforderlichkeit der Datenverarbeitung als auch der Forderung nach angemessenen Schutzmaßnahmen weitestmöglich Rechnung und bietet somit die beste Gewähr dafür, dass eine Interessenabwägung zugunsten der Forschung ausfallen wird. Die Datentreuhand ist damit bei einem solchen Verständnis zwar keine unverzichtbare Voraussetzung für eine zulässige Forschungsdatenverarbeitung, kann sich aber zu einem „Goldstandard" der Forschungsdatenverarbeitung entwickeln.

4 Ausblick: Die Datentreuhand im 21. Jahrhundert

Die Zurverfügungstellung von Daten für Forscher in anonymer Form gehört zu den wichtigsten Aufgaben einer Datentreuhand. Nationale und internationale Statistikinstitute waren Vorreiter bei der Anonymisierung von Daten durch Modifikation (beispielsweise wurde das Konzept der Anonymisierung durch Löschung in diesem Bereich bereits 1977 formalisiert[26]). Durch den digitalen Wandel entstehen jedoch zunehmend mehr und größere Datenbestände, die durch Verknüpfung dazu genutzt werden können, Datensubjekte relativ einfach auch in modifizierten Datenbeständen zu reidentifizieren[27] (der wohl spektakulärste Fall war die Reidentifikation der Patientenakte eines amerikanischen Politikers[28]). Viele Experten gehen deshalb davon aus, dass heutzutage durch klassische Anonymisierungsverfahren keine ausreichende Anonymität mehr gewährleistet werden kann.[29] Um ihren Auftrag erfüllen zu können, muss die Datentreuhand im 21. Jahrhundert deshalb auf moderne technische Verfahren zur Anonymisierung setzen, die teilweise noch Gegenstand laufender Forschung sind.

- Eine innovative Technik zur Wahrung der Anonymität ist die sogenannte *Differential Privacy*, die einen Zufallsfaktor bei der Verarbeitung von Daten (bspw. im Rahmen der Herausgabe durch die Treuhand oder einer bei der Treuhand durchgeführten Auswertung) voraussetzt und, wenn sie richtig parametrisiert ist, sehr starke Garantien für den Schutz der Privatsphäre geben kann. Auch wenn das 2006 erstmals vorgestellte Verfahren[30] mittlerweile sogar von den großen Internet-Konzernen, wie Apple, Google und Facebook, zunehmend in der Praxis eingesetzt wird, gibt es bzgl. eines Einsatzes in der Medizin Bedenken wegen des Nichtdeterminismus aufgrund des Zufallsfaktors.
- Mittels moderner Verfahren der künstlichen Intelligenz können heutzutage auch „künstliche" bzw. *synthetische* Daten nach dem Muster eines Originaldatensatzes erzeugt werden, die vergleichbare statistische Eigenschaften haben. Dieser Prozess kann mit Differential Privacy kombiniert werden, um einen garantierten Schutz der Privatsphäre der im Originaldatensatz abgebildeten Subjekte zu erreichen.
- Beim Konzept der *Datenenklave* werden die Daten vom Treuhänder verwaltet – aber nicht herausgegeben. Stattdessen werden virtuelle Arbeitsumgebungen angeboten, in denen berechtigte Forscher die Daten unter (virtueller, automatisierter) Aufsicht verarbeiten können. Das Konzept wird in Großbritannien bereits heute erfolgreich für die Zusammenführung und die Sekundärnutzung von Gesundheitsdaten eingesetzt. Eine solche Datenenklave könnte auch für die im FDZ zu speichernden Daten erwogen werden, um die Problematik des Reidentifikationsrisikos bei mehrfacher Herausgabe unterschiedlicher Ausschnitte des anonymisierten Datenbestands zu vermeiden.
- Umgesetzt werden kann das Prinzip der Datenenklave auch unter Einsatz moderner kryptographischer Verfahren, die eine Treuhand für verteilt vorliegende Daten verlässlich simulieren können. Bei der sogenannten *Secure Multiparty Computation* und verwandten Verfahren wie der homomorphen Verschlüsselung können mathematische Operationen auf verschlüsselten Daten ausgeführt werden.

Mit einem Bündel aus solchen innovativen Technologien kann eine Treuhand heute und auch in der Zukunft einen starken Schutz der Privatsphäre der betroffenen Personen gewährleisten.

Open Access

25 Zur Herausforderung, dass sich gerade im Falle komplexer klinischer Daten ab einem gewissen Punkt immer ein Reidentifizierungsrisiko stellt (wenn auch mit normalerweise unverhältnismäßigem Aufwand), s. sogleich unten 4.

26 *Dalenius*, Towards a methodology for statistical disclosure control. Stat Tidskr, 1977, 15: S. 429-444.

27 *Rocher/Hendrickx/De Montjoye*, Estimating the success of re-identifications in incomplete datasets using generative models. Nature communications, 2019 Jul 23;10(1):1-9.

28 *Sweeney, L.*, k-Anonymity: A model for protecting privacy. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2002, 10(05): S. 557-570.

29 *Ohm*, Broken promises of privacy: Responding to the surprising failure of anonymization, UCLA L. Rev. 2009; 57:1701.

30 *Dwork/McSherry/Nissim/Smith*, Calibrating noise to sensitivity in private data analysis. In: Theory of cryptography conference 2006 Mar 4 (S. 265-284).