Herausforderungen an das Datenschutzrecht durch den Einsatz von Künstlicher Intelligenz in der Medizin

Tobias Herbst

I. Einleitung

Der Einsatz von Künstlicher Intelligenz (KI) in der Medizin wirft einige datenschutzrechtliche Fragen und Probleme auf. Je nach konkretem Anwendungsbereich betrifft das insbesondere die Intransparenz des Zustandekommens von Ergebnissen oder den Umfang personenbezogener Daten, die für manche KI-Anwendungen erforderlich sind. Im Folgenden sollen zunächst einige wesentliche Grundsätze des geltenden Datenschutzrechts erläutert werden, die durch den Einsatz von KI in der Medizin berührt werden können. Danach werden Anwendungsbeispiele aus den Bereichen der medizinischen Versorgung und der medizinischen Forschung einschließlich möglicher Kollisionen mit dem geltenden Datenschutzrecht (und deren Vermeidung) erörtert. Ein Fazit beschließt den Beitrag.

II. Wesentliche Grundsätze des geltenden Datenschutzrechts

In der Datenschutzgrundverordnung der EU (DSGVO)[1] fasst Art. 5 „Grundsätze für die Verarbeitung personenbezogener Daten" zusammen; Art. 9 stellt darüber hinaus für sogenannte „besondere Kategorien personenbezogener Daten", zu denen auch für die Medizin relevante Daten zählen können, zusätzliche Anforderungen auf.

1 Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung), ABl. EU Nr. L 119, S. 1.

I. Die Datenschutzgrundsätze des Art. 5 DSGVO

Im Folgenden sollen zunächst die in Art. 5 Abs. 1 DSGVO niedergelegten Datenschutzgrundsätze[2] im Hinblick auf ihre Relevanz für den Einsatz von KI in der Medizin betrachtet werden. Es handelt sich um die Grundsätze der Rechtmäßigkeit, Fairness, Transparenz, Zweckbindung, Datenminimierung, Richtigkeit, Speicherbegrenzung sowie Integrität und Vertraulichkeit.

a) Rechtmäßigkeit

Personenbezogene Daten müssen „auf rechtmäßige Weise" verarbeitet werden. Damit ist gemeint, dass es für die Verarbeitung eine Rechtsgrundlage im Unionsrecht oder im nationalen Recht geben muss, also eine Rechtsnorm, die die konkrete Datenverarbeitung erlaubt.[3] So erlaubt etwa Art. 6 Abs. 1 lit. a DSGVO die Datenverarbeitung, wenn die betroffene Person ihre Einwilligung dazu gegeben hat; werden „besondere Kategorien personenbezogener Daten" verarbeitet, dann muss sich die Einwilligung gemäß Art. 9 Abs. 2 lit. a DSGVO außerdem ausdrücklich hierauf beziehen. Für vertraglich vereinbarte Datenverarbeitungen (z.B. im Rahmen eines Behandlungsvertrages) enthält Art. 6 Abs. 1 lit. b DSGVO eine Rechtsgrundlage. Im medizinischen Kontext jedenfalls theoretisch denkbar ist auch Art. 6 Abs. 1 lit. d DSGVO als Rechtsgrundlage, wonach die Datenverarbeitung zum Schutz lebenswichtiger Interessen erlaubt ist. Für Datenverarbeitungen, die der Erfüllung rechtlicher Verpflichtungen (wie etwa Dokumentationspflichten bei medizinischen Behandlungen) oder der Wahrnehmung von Aufgaben im öffentlichen Interesse dienen (Letzteres trifft in der Regel für medizinische Forschung zu), eröffnen die Regelungen in Art. 6 Abs. 1 lit. c und e in Verbindung mit Art. 6 Abs. 3 DSGVO die Möglichkeit zur Schaffung von Rechtsgrundlagen im nationalen Recht.[4] Insbesondere für KI-basierte Forschung mit großen Mengen von Daten aus dem Behandlungszusammenhang kann das Bestehen einer ausreichenden Rechtsgrundlage fraglich sein; darauf wird bei den Anwendungsbeispielen zurückzukommen sein.

2 Vgl. dazu auch *T. Herbst*, in: J. Kühling/B. Buchner (Hrsg.), DS-GVO/BDSG, 4. Aufl., München 2024, Art. 5 Rn. 7 ff.

3 *S. Pötters*, in: P. Gola/D. Heckmann, Datenschutz-Grundverordnung/Bundesdatenschutzgesetz, 3. Aufl., München 2022, Art. 5 Rn. 7.

4 *B. Buchner/T. Petri*, in: Kühling/Buchner (Hrsg.), DS-GVO/BDSG (Fn. 2), Art. 6 Rn. 83 ff., 120 ff.

b) Fairness

Der Grundsatz der Fairness („Verarbeitung nach Treu und Glauben") verlangt im Wesentlichen, dass die betroffene Person bei der Verarbeitung ihrer Daten fair behandelt wird, dass also eine Art Kräftegleichgewicht besteht zwischen ihr und demjenigen, der ihre Daten verarbeitet (dem „Verantwortlichen").[5] Ein Verstoß gegen diesen Grundsatz kann im medizinischen Kontext etwa dann vorliegen, wenn ein Forscher bzw. ein behandelnder Arzt seinen Wissensvorsprung bzw. den Umstand, dass ein Patient auf ihn angewiesen ist, dazu ausnutzt, um (etwa für Forschungszwecke) an Daten zu gelangen. Auch eine Anwendung von KI, die zu diskriminierenden Ergebnissen führt, kann als Verstoß gegen den Grundsatz der Fairness angesehen werden.[6]

c) Transparenz

Nach dem Grundsatz der Transparenz müssen personenbezogene Daten in einer für die betroffene Person nachvollziehbaren Weise verarbeitet werden. Das bedeutet insbesondere, dass die betroffene Person umfassend über die Verarbeitung ihrer Daten zu informieren ist.[7] Im Kontext der Anwendung von KI in der Medizin stellt sich in diesem Zusammenhang die Frage, wie weit die Funktionsweise der KI oder gar das Zustandekommen bestimmter Ergebnisse dem Betroffenen erläutert werden müssen. Darauf wird bei den Anwendungsbeispielen zurückzukommen sein.

d) Zweckbindung

Der Zweckbindungsgrundsatz verlangt, dass bei der Erhebung von Daten der Zweck ihrer Verarbeitung festgelegt wird und dass im Weiteren diese Daten nur entweder zu diesem Zweck oder – bei einer Zweckänderung – zu einem anderen Zweck, der mit dem Erhebungszweck vereinbar ist, verarbeitet werden dürfen.[8] Im Bereich der KI-basierten medizinischen Forschung liegt eine Zweckänderung etwa dann vor, wenn Daten, die

5 Vgl. *P. Reimer*, in: G. Sydow/N. Marsch (Hrsg.), DS-GVO/BDSG, 3. Aufl., Baden-Baden 2022, Art. 5 Rn. 14 f.
6 So etwa *P. Vogel*, Künstliche Intelligenz und Datenschutz, Baden-Baden 2022, S. 105.
7 Vgl. *Reimer* (Fn. 5), Art. 5 Rn. 16 f.
8 Dazu ausführlich m.w.N. *Herbst* (Fn. 2), Art. 5 Rn. 20 ff.

zunächst im Behandlungskontext und damit zum Zweck der Behandlung erhoben wurden, im Nachhinein mit KI-Methoden zu Forschungszwecken ausgewertet werden. Eine solche sogenannte Sekundärnutzung solcher Daten ist nur zulässig, wenn der Forschungszweck mit dem ursprünglichen Erhebungszweck vereinbar ist. In diesem Zusammenhang sieht Art. 5 Abs. 1 lit. b DSGVO eine Privilegierung u.a. der Forschung vor: Eine Weiterverarbeitung für wissenschaftliche Forschungszwecke „gemäß Artikel 89 Absatz 1" gilt als vereinbar mit den ursprünglichen Zwecken. Der Verweis auf Art. 89 Abs. 1 DSGVO hat dabei zur Konsequenz, dass nur solche Datenverarbeitungen für Forschungszwecke als vereinbar mit dem ursprünglichen Erhebungszweck gelten, bei denen besondere Vorkehrungen in Gestalt technischer und organisatorischer Maßnahmen getroffen werden; das bedeutet insbesondere, dass die entsprechenden Daten soweit möglich pseudonymisiert oder anonymisiert werden müssen. Werden diese Vorgaben beachtet, dann kann eine Sekundärnutzung medizinischer Daten zu Forschungszwecken dem Zweckbindungsgrundsatz genügen. Allerdings ergibt sich hier noch ein weiteres Problem, denn für die Verarbeitung zu Forschungszwecken ist auch eine geeignete Rechtsgrundlage erforderlich; ein Behandlungsvertrag oder eine Einwilligung in die Datenverarbeitung zu Behandlungszwecken genügt für sich genommen hierfür in der Regel nicht.[9] Als mögliche Rechtsgrundlagen kommen hier Forschungseinwilligungen und besondere „Forschungsklauseln" im nationalen Recht in Betracht. Auf die damit zusammenhängenden Fragen soll bei den Anwendungsbeispielen eingegangen werden.

e) Datenminimierung

Der Grundsatz der Datenminimierung beschränkt die Datenverarbeitung und damit den Umfang der gespeicherten und verarbeiteten personenbezogenen Daten auf das für die jeweils festgelegten Zwecke notwendige Maß.[10]

9 In der Literatur wird auch die Meinung vertreten, dass bei einer Zweckänderung, die mit dem Erhebungszweck vereinbar ist, die Rechtsgrundlage der ursprünglichen Datenerhebung ausreicht. Ausführlich zu dieser Kontroverse *Herbst* (Fn. 2), Art. 5 Rn. 48 ff. Die Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder (DSK) hat sich inzwischen gegen die Ansicht gewandt, dass hier die Rechtsgrundlage der Datenerhebung ausreiche: *DSK*, Erfahrungsbericht der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder zur Anwendung der DS-GVO, November 2019, S. 13 f.

10 Dazu etwa *Reimer* (Fn. 5), Art. 5 Rn. 32 ff.

Im Unterschied zur Regelung im „alten" Datenschutzrecht (§ 3a BDSG a.F.), das allgemein ein Erfordernis der Datenvermeidung und Datensparsamkeit aufstellte, stellt der Grundsatz der Datenminimierung eine Relation zwischen dem Umfang der Datenverarbeitung und ihrem Zweck her. Für die KI-basierte Auswertung großer Mengen medizinischer Daten bedeutet das, dass sie nun nicht mehr von vornherein als datenschutzrechtlich problematisch anzusehen ist – soweit sich der Zweck der Verarbeitung etwa im Bereich der medizinischen Forschung nur mithilfe der Auswertung großer Datenmengen erreichen lässt, ist dies mit dem Grundsatz der Datenminimierung vereinbar. Ein Problem mit diesem Grundsatz kann sich aber im Fall der Sekundärnutzung ergeben, also dann, wenn medizinische Daten zunächst in einem reinen Behandlungskontext anfallen und erst später für Forschungszwecke verwendet werden sollen. In dieser Situation kann der Grundsatz der Datenminimierung dazu führen, dass Daten, die in einem Forschungskontext wertvoll sein könnten, nicht erhoben bzw. verarbeitet werden (dürfen), weil sie für den Behandlungskontext nicht notwendig sind. Das verringert die Menge der für die Forschung potentiell zur Verfügung stehenden Daten.

f) Richtigkeit

Nach dem Grundsatz der Richtigkeit müssen personenbezogene Daten „sachlich richtig und erforderlichenfalls auf dem neuesten Stand sein". Bei der Anwendung von Techniken des Machine Learning mithilfe neuronaler Netze könnte man hier die Schwierigkeit vermuten, medizinische Daten auf den neuesten Stand zu bringen; nachträgliche Änderungen an einem neuronalen Netz im Sinne einer Aktualisierung einzelner Datensätze, die zu dessen Training verwendet wurden, sind nicht ohne Weiteres möglich. Allerdings gilt das Erfordernis der Aktualität der Daten nur im Hinblick auf die jeweiligen Zwecke ihrer Verarbeitung.[11] Im Bereich der medizinischen Forschung etwa dürfte es für die Erreichung der Zwecke der Verarbeitung in der Regel nicht darauf ankommen, ob es sich um aktuelle oder historische Daten handelt. Der Grundsatz der Richtigkeit verlangt dann nicht nach einer „Aktualisierung" eines neuronalen Netzes.

11 *H. Heberlein*, in: E. Ehmann/M. Selmayr (Hrsg.), DS-GVO, 2. Aufl., München 2018, Art. 5 Rn. 24.

g) Speicherbegrenzung

Der Grundsatz der Speicherbegrenzung betrifft die zeitliche Dauer der Verarbeitung personenbezogener Daten: Diese dürfen nur so lange verarbeitet werden, wie es für die Zwecke, für die sie verarbeitet werden, erforderlich ist. Wird allerdings der Personenbezug aufgehoben, indem die Daten anonymisiert werden, dann darf die Verarbeitung über diesen Zeitpunkt hinaus fortgesetzt werden; das ergibt sich schon daraus, dass Daten ohne Personenbezug ohnehin aus dem Anwendungsbereich der DSGVO herausfallen, wird aber auch im Wortlaut des Art. 5 Abs. 1 lit. e DSGVO zum Ausdruck gebracht (eine Identifizierung darf „nur so lange ermöglicht [werden], wie es für die Zwecke, für die sie verarbeitet werden, erforderlich ist"). Eine Auswertung von Daten etwa aus dem medizinischen Behandlungskontext durch KI-basierte Methoden ist also nach einer Anonymisierung dieser Daten auch dann zulässig, wenn die Daten für den ursprünglichen Erhebungszweck (Behandlung) nicht mehr benötigt werden.[12] Bei der Anonymisierung ist aber die Gefahr einer Re-Identifizierung zu beachten, die gerade bei der Auswertung großer Datenmengen durch KI-basierte Methoden besteht; dieser Gefahr muss dann ggf. etwa durch Reduzierung des Informationsgehaltes, z.B. durch die Umwandlung in aggregierte Daten, begegnet werden.

Ähnlich wie beim Grundsatz der Zweckbindung normiert die DSGVO auch bei der Speicherbegrenzung eine Privilegierung u.a. der Forschung: Bei Beachtung der Anforderungen des Art. 89 Abs. 1 DSGVO (also Durchführung technischer und organisatorischer Maßnahmen wie Pseudonymisierung oder Anonymisierung) dürfen Daten auch länger als etwa für ein konkretes Forschungsprojekt erforderlich verarbeitet (und damit gespeichert) werden, so dass sie später auch für weitere Forschungsprojekte zur Verfügung stehen können. Dennoch kann der Grundsatz der Speicherbegrenzung eine Einschränkung der Möglichkeiten der Sekundärnutzung von Behandlungsdaten für die Forschung bewirken; denn sobald solche Daten für den Behandlungskontext nicht mehr erforderlich sind, dürfen sie nicht mehr verarbeitet werden, müssen also gelöscht oder anonymisiert werden. Die Überführung von Behandlungsdaten in pseudonymisierte Forschungsdaten ist daher nur in dem Zeitraum möglich, in dem diese Daten noch für die Behandlung benötigt werden.

12 *A. Roßnagel*, in: S. Simitis/G. Hornung/I. Spiecker gen. Döhmann (Hrsg.), Datenschutzrecht, Baden-Baden 2019, Art. 5 Rn. 150 ff.

h) Integrität und Vertraulichkeit

Der Grundsatz der Integrität und Vertraulichkeit verlangt, dass personenbezogene Daten in einer Weise verarbeitet werden, die eine angemessene Sicherheit dieser Daten gewährleistet; das betrifft insbesondere (aber nicht nur) den Schutz vor unbefugtem Zugriff Dritter, also z.B. vor Hackerangriffen.[13] Im Bereich KI-basierter Forschung ist dabei insbesondere zu beachten, dass gerade große Datenmengen besonders attraktiv für solche Angriffe sein können.

2. Besondere Anforderungen an medizinische Daten

Die Regelungen in Art. 9 DSGVO stellen besondere Anforderungen an die Verarbeitung bestimmter besonders „sensibler" Daten – die DSGVO spricht hier von „besonderen Kategorien personenbezogener Daten". Zu diesen Datenarten zählen unter anderem genetische Daten, biometrische Daten zur eindeutigen Identifizierung einer natürlichen Person, Gesundheitsdaten sowie Daten zum Sexualleben und zur sexuellen Orientierung.[14] Medizinisch relevante Daten, also insbesondere Daten, die im Behandlungskontext anfallen, dürften in der Regel jeweils mindestens einer dieser Kategorien angehören. Die Regelungen in Art. 9 DSGVO sind in der Weise strukturiert, dass Art. 9 Abs. 1 zunächst die Verarbeitung solcher Daten generell untersagt, während Art. 9 Abs. 2 dieses allgemeine Verarbeitungsverbot für bestimmte Fallgruppen unter bestimmten Voraussetzungen wieder aufhebt. Für die Thematik dieses Beitrags ist dabei zunächst Art. 9 Abs. 2 lit. a DSGVO relevant, wonach die Verarbeitung entsprechender sensibler Daten zulässig ist, wenn die betroffene Person hierin ausdrücklich einwilligt. Eine solche ausdrückliche Einwilligung hat etwa die Gestalt: „Ich bin auch mit der Verarbeitung von Daten über meine Gesundheit einverstanden".

Außerdem eröffnet Art. 9 Abs. 2 lit. j DSGVO den Mitgliedstaaten der EU die Möglichkeit, in ihrem nationalen Recht die Verarbeitung entsprechender sensibler Daten für Forschungszwecke vorzusehen; dabei müssen aber der Verhältnismäßigkeitsgrundsatz und der Wesensgehalt des Rechts

13 Vgl. *Reimer* (Fn. 5), Art. 5 Rn. 48 ff.

14 Dazu und zum Folgenden *T. Petri*, in: Simitis/Hornung/Spiecker gen. Döhmann (Hrsg.), Datenschutzrecht (Fn. 12), Art. 9 Rn. 10 ff.

auf Datenschutz gewahrt und Maßnahmen zur Wahrung der Grundrechte und Interessen der betroffenen Person getroffen werden wie etwa technische und organisatorische Maßnahmen nach Art. 89 Abs. 1 DSGVO.[15] Der Bundesgesetzgeber und die verschiedenen Landesgesetzgeber haben in Deutschland von dieser Möglichkeit in verschiedenen sogenannten „Forschungsklauseln" Gebrauch gemacht (so auf Bundesebene in § 27 BDSG); darauf wird noch bei den Anwendungsbeispielen zurückzukommen sein.

III. Anwendungsbereiche

Beispiele für den Einsatz von KI in der Medizin finden sich in den beiden Bereichen der medizinischen Versorgung und der medizinischen Forschung.

1. Medizinische Versorgung

Als Beispiel für den Einsatz von KI in der medizinischen Versorgung soll im Folgenden eine Gesundheits-App dienen, die Funktionen der Diagnose und Therapie im Bereich der Psychiatrie erfüllen soll. Die App kommuniziert mit dem Anwender mithilfe von Text oder auch gesprochener Sprache und erstellt KI-basiert Diagnosen und Therapieanweisungen (etwa mit Methoden der Verhaltenstherapie). Abgesehen von rechtlichen Problemen, die sich daraus ergeben, dass eine solche App ein Medizinprodukt im Sinne der Medizinprodukteverordnung der EU darstellen kann,[16] stellt auch das Datenschutzrecht Anforderungen an eine solche App.[17]

Solche Anforderungen ergeben sich zunächst aus Art. 22 DSGVO. Diese Vorschrift regelt die Zulässigkeit sogenannter automatisierter Entscheidungen im Einzelfall. Darunter versteht die Regelung eine vollständig automatisiert (also ohne Mitwirkung eines Menschen im Einzelfall) getroffene Entscheidung, die gegenüber einer Person rechtliche Wirkung entfaltet oder sie in ähnlicher Weise erheblich beeinträchtigt.[18] Sofern die eben

15 *Petri* (Fn. 14), Art. 9 Rn. 98.

16 Ausführlich dazu *Z. Schreitmüller*, Regulierung intelligenter Medizinprodukte, Baden-Baden 2023, S. 109 ff.

17 Dazu wiederum *Schreitmüller* (Fn. 16), S. 165 ff.

18 Dazu etwa *K. v. Lewinski*, in: H. A. Wolff/S. Brink/A. v. Ungern-Sternberg (Hrsg.), BeckOK Datenschutzrecht, 44. Edition, München 2023, Art. 22 Rn. 14 ff.

dargestellte App Therapieanweisungen vollautomatisiert erstellt, stellen diese Anweisungen möglicherweise automatisierte Entscheidungen im Sinne dieser Vorschrift dar, denn die Therapieanweisungen können – aus welchen Gründen auch immer – anstelle der gewünschten Genesungseffekte auch gesundheitliche Beeinträchtigungen nach sich ziehen und damit den Anwender erheblich beeinträchtigen.

Die Regelungen in Art. 22 DSGVO sind derart strukturiert, dass Art. 22 Abs. 1 zunächst ein generelles Verbot solcher automatisierter Entscheidungen normiert, während Art. 22 Abs. 2 einige Ausnahmen von diesem Verbot regelt. Aufgrund des Umstandes, dass die oben dargestellte App Gesundheitsdaten und damit besondere Kategorien von Daten im Sinne des Art. 9 DSGVO verarbeitet, kommt von diesen Ausnahmen letztlich nur die ausdrückliche Einwilligung der betroffenen Person in Betracht (vgl. Art. 22 Abs. 4 DSGVO mit dem Verweis auf Art. 9 DSGVO). In dem Beispiel müsste also der Anwender der App ausdrücklich darin einwilligen, dass Diagnose und Therapieanweisungen vollständig automatisiert und ohne menschliches Zutun erstellt werden.

Darüber hinaus verlangt Art. 13 Abs. 2 lit. f DSGVO, dass die betroffene Person nicht nur die allgemein bei der Verarbeitung personenbezogener Daten verpflichtenden Informationen erhält, sondern auch spezifische Informationen über die automatisierte Entscheidung im Einzelfall. Die Regelung verlangt dabei, dass der Anwender nicht nur über das Bestehen einer automatisierten Entscheidungsfindung informiert wird, sondern auch „aussagekräftige Informationen über die involvierte Logik sowie die Tragweite und die angestrebten Auswirkungen einer derartigen Verarbeitung" erhält. Während die Aufklärung über die Tragweite und die Auswirkungen der Anwendung der App sich vielleicht noch im Rahmen der ärztlichen Aufklärung im Vorfeld etwa einer Verhaltenstherapie hält, stellt sich die Frage, welche aussagekräftigen Informationen „über die involvierte Logik" bezüglich der eingesetzten KI gegeben werden sollen oder können.

Schon die Verwendung des Ausdrucks „Logik" in dieser Vorschrift deutet darauf hin, dass sich der Gesetzgeber beim Erlass der DSGVO hier eher eine Entscheidungsfindung mithilfe von herkömmlichen Algorithmen im Sinne von klar definierten und vorprogrammierten Rechenschritten vorstellte. Allerdings stellt sich schon bei solchen Algorithmen – jedenfalls ab einer gewissen Komplexität – die Frage, wie sie einem Betroffenen zu erklären wären. Unter der Geltung des früheren Bundesdatenschutzgesetzes (BDSG) hatte der BGH zur Information des Betroffenen über die

Methoden der Berechnung der Kreditausfallwahrscheinlichkeit durch die Schufa geurteilt, dass der Betroffene lediglich einen Anspruch auf Auskunft darüber hat, welche kreditrelevanten Daten in die Berechnung des Wahrscheinlichkeitswertes eingeflossen sind; einen Anspruch auf Information über die abstrakte Methode der Wahrscheinlichkeitsberechnung (Scorewertberechnung) hat der BGH ausdrücklich verneint.[19] Die entsprechende Regelung im früheren BDSG normierte einen Auskunftsanspruch über „das Zustandekommen und die Bedeutung der Wahrscheinlichkeitswerte einzelfallbezogen und nachvollziehbar in allgemein verständlicher Form" (§ 34 Abs. 4 Nr. 4 BDSG a.F.). Bei der Entscheidung des BGH spielte vor allem eine Rolle, dass die Berechnungsmethode aus Sicht der Schufa ein Geschäftsgeheimnis darstellte; in erster Linie ging es also nicht um die Möglichkeit oder Unmöglichkeit der Information des Betroffenen über komplexe Algorithmen.

Jenseits der hier nicht zu beantwortenden Frage, welche Anforderungen die aktuelle Regelung der DSGVO hinsichtlich eines Auskunftsanspruchs des Betroffenen gegenüber der Schufa stellt, ist zu fragen, in welcher Weise nach Art. 13 Abs. 2 lit. f DSGVO über die „involvierte Logik" einer KI aufzuklären ist. Wird etwa ein mit bestimmten Daten trainiertes neuronales Netz verwendet, dann ist es heute selbst für Hersteller und Betreiber eines solchen neuronalen Netzes häufig nicht möglich, das Zustandekommen der verschiedenen Gewichtungen innerhalb des Netzes nachzuvollziehen und zu erklären, auch wenn es hierfür in einzelnen Bereichen Ansätze gibt („Erklärbare KI" oder „Explainable AI").[20] Allerdings genügt es dabei nicht, wie häufig bei Methoden der Erklärbaren KI, im Nachhinein, also nach Vollendung des Verarbeitungsvorganges, das Zustandekommen des Ergebnisses zu erklären, denn Art. 13 Abs. 2 DSGVO verlangt ausdrücklich, dass die betreffenden Informationen der betroffenen Person „zum Zeitpunkt der Erhebung dieser Daten" zur Verfügung gestellt werden müssen. Letztlich wird man sich bei der Information über die KI auf das beschränken müssen, was sinnvoll und möglich ist. Dabei muss aber berücksichtigt werden, welche Informationen dennoch für den Betroffenen relevant sein können.[21] Das können z.B. die fehlende exakte Vorhersehbarkeit der Ergebnisse bzw. die Wahrscheinlichkeit „richtiger" oder „guter" Ergebnisse, die Herkunft

19 BGHZ 200, 38 (42 ff.).

20 Vgl. dazu im Überblick *C. Niederée/W. Nejdl*, in: M. Ebers/C. Heinze/T. Krügel/B. Steinrötter (Hrsg.), Künstliche Intelligenz und Robotik, München 2020, § 2 Rn. 123 ff.

21 Ähnlich *Schreitmüller*, Regulierung (Fn. 16), S. 249 f.; *Vogel*, Künstliche Intelligenz (Fn. 6), S. 172 ff.

der Trainingsdaten eines neuronalen Netzes, mögliche Parallelen zwischen den Ergebnissen und den Trainingsdaten (einschließlich eines möglichen schon in den Trainingsdaten vorliegenden, den Betroffenen benachteiligenden Bias) oder etwaige Beschränkungen hinsichtlich der möglichen Ergebnisse (im Beispiel etwa ein „einprogrammierter" Ausschluss bestimmter Therapieanweisungen) sein.

2. Medizinische Forschung

Ein weiterer Anwendungsbereich des Einsatzes von KI in der Medizin ist die medizinische Forschung. Typischerweise geht es hier um die Auswertung großer Datenmengen mit KI-basierten Methoden. Im Folgenden sollen zunächst einige Beispiele aktueller Forschungsprojekte mit großen Datenmengen vorgestellt werden, bei denen auch KI zur Anwendung kommen kann. In datenschutzrechtlicher Hinsicht stellen sich hier vor allem Probleme im Hinblick auf geeignete Rechtsgrundlagen für die Verarbeitung der benötigten großen Datenmengen. Als eine solche Rechtsgrundlage kommt in erster Linie die Einwilligung der Betroffenen in Betracht; entsprechend ist dann auf die einwilligungsbasierte Forschung mit entsprechenden Daten einzugehen. Soweit es an einer Einwilligung der Betroffenen in die Verarbeitung ihrer Daten für Forschungszwecke fehlt, bedarf es für solche Verarbeitungen einer speziellen gesetzlichen Grundlage. Welche gesetzlichen Grundlagen hierfür in Betracht kommen, soll im Anschluss erörtert werden. Schließlich sollen noch einige datenschutzrechtliche Fragen erörtert werden, die sich bei den für die Forschung wichtigen internationalen Kooperationen ergeben können.

a) Beispiele für potentiell KI-basierte medizinische Forschung mit großen Datenmengen

In der vom Bundesministerium für Bildung und Forschung geförderten Medizininformatik-Initiative (MII) haben sich Wissenschaftler aus den deutschen Universitätskliniken zusammengeschlossen, um Patientendaten,

die während eines Klinikaufenthaltes im Behandlungskontext entstehen, digital zu vernetzen, so dass mit diesen Daten geforscht werden kann.[22]

Ein anderes Projekt ist die von einem Netzwerk deutscher Forschungseinrichtungen durchgeführte NAKO Gesundheitsstudie. Es handelt sich dabei um eine auf eine Dauer von 20 bis 30 Jahren angelegte Langzeit-Bevölkerungsstudie zur Erforschung der Ursachen für die Entstehung von Volkskrankheiten. Hierzu werden ca. 200.000 zufällig ausgewählte Teilnehmer umfassend und über einen längeren Zeitraum wiederkehrend medizinisch untersucht.[23]

Einen weiteren Forschungskontext mit ggf. großen Datenmengen bilden die insbesondere an vielen Universitätskliniken (in Deutschland und weltweit) eingerichteten Biobanken. Hier werden menschliche Biomaterialproben auf Dauer eingelagert und zugehörige medizinische Daten gespeichert. Proben und Daten können dann für Forschungszwecke unter bestimmten Voraussetzungen zur Verfügung gestellt werden.[24]

b) Einwilligung als Grundlage der KI-basierten Datenverarbeitung

Bei den soeben vorgestellten Projekten der MII und der NAKO beruht die Verarbeitung medizinischer Daten zu Forschungszwecken auf den Einwilligungen der Patienten bzw. Teilnehmer in die Forschung mit ihren Daten. In Biobanken wird ein Teil der Proben und Daten zwar auch im reinen Behandlungskontext aufbewahrt (etwa um spätere Nach- oder Vergleichsuntersuchungen in Bezug auf den betreffenden Patienten durchführen zu können); in den letzten Jahren sind aber jedenfalls die größeren Biobanken bestrebt, auch Forschungseinwilligungen zu erhalten, um auf dieser Grundlage Proben und Daten für künftige Forschungsprojekte zur Verfügung stellen zu können.

Charakteristisch für solche Forschung mit großen Datenmengen ist, dass die mit großem Aufwand gewonnenen Daten nicht nur für ein einzelnes bestimmtes Forschungsthema verwendet werden sollen, sondern über einen längeren Zeitraum für verschiedene zukünftige, zum Zeitpunkt der Datenerhebung oft noch nicht bekannte Forschungsprojekte zur Verfügung

22 Vgl. dazu die Internetpräsenz der MII unter https://www.medizininformatik-initiative.de.

23 Vgl. dazu die Internetpräsenz der NAKO unter https://nako.de.

24 Zur Biobank-Forschung vgl. etwa die Internetpräsenz des German Biobank Node https://www.bbmri.de.

stehen sollen. Für einwilligungsbasierte Forschung bedeutet dies, dass die Einwilligung der betroffenen Personen auch solche erst in der Zukunft liegenden Forschungen abdecken muss. Schon aus praktischen Gründen ist es bei großen Datenmengen und langen Zeiträumen oft nicht möglich, für jedes künftige Forschungsprojekt bei den betroffenen Personen nachzufragen und eine erneute Einwilligung einzuholen. Daher wird in diesen Konstellationen angestrebt, schon bei Gewinnung der Daten eine möglichst umfassende Einwilligung zu erlangen, die von vornherein ein breites Spektrum künftiger Forschungsprojekte abdeckt; man spricht hier von einem „broad consent". Eine solche Einwilligung erstreckt sich dann z.B. auf jegliche medizinische Forschung, die die Vorbeugung, Erkennung und Behandlung von Erkrankungen verbessern soll.

Die Zulässigkeit eines solchen broad consent wird manchmal angezweifelt.[25] In datenschutzrechtlicher Hinsicht wird hier etwa verwiesen auf die Legaldefinition der Einwilligung in Art. 4 Nr. 11 DSGVO, wonach zu den Merkmalen einer Einwilligung zählt, dass sie eine Willensbekundung ist, die „für den bestimmten Fall" abgegeben wird. Diese Formulierung wird allerdings durch den auf Zwecke der wissenschaftlichen Forschung bezogenen Erwägungsgrund 33 zur DSGVO relativiert. Dort heißt es:

„Oftmals kann der Zweck der Verarbeitung personenbezogener Daten für Zwecke der wissenschaftlichen Forschung zum Zeitpunkt der Erhebung der personenbezogenen Daten nicht vollständig angegeben werden. Daher sollte es betroffenen Personen erlaubt sein, ihre Einwilligung für bestimmte Bereiche wissenschaftlicher Forschung zu geben, wenn dies unter Einhaltung der anerkannten ethischen Standards der wissenschaftlichen Forschung geschieht."

Die Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder („Datenschutzkonferenz") hat im Hinblick auf diesen Erwägungsgrund in einem Beschluss vom 3.4.2019 festgestellt, dass der Ansatz des broad consent zum Tragen kommen kann, wenn das Forschungsvorhaben eine vollständige Zweckbestimmung schlechthin nicht zulässt und wenn bestimmte Vorkehrungen getroffen werden wie etwa ein positives Votum eines Ethikgremiums vor der Nutzung für weitere

25 So bezüglich der von der MII verwendeten Mustereinwilligung von *W. Fröhlich/I. Spiecker gen. Döhmann*, Die breite Einwilligung (Broad Consent) in die Datenverarbeitung zu medizinischen Forschungszwecken – der aktuelle Irrweg der MII, GesR 2022, 346 (349 ff.).

Forschungszwecke oder die Einrichtung einer Internetpräsenz, über die sich Teilnehmer über laufende und künftige Studien informieren können.[26]

Bei Beachtung solcher Vorkehrungen ist demnach ein broad consent datenschutzrechtlich zulässig. Selbstverständlich stellt der broad consent nur dann eine wirksame Einwilligung dar, wenn die betroffenen Personen zuvor klar und unmissverständlich über die Breite der Einwilligung, also das Spektrum möglicher Forschungen, informiert wurden und wenn die späteren Forschungsprojekte sich innerhalb des so abgesteckten Rahmens halten. Soweit es potentiellen Studienteilnehmern auf diese Weise ermöglicht wird, ihre Daten (und ggf. Biomaterialproben) bewusst für medizinische Forschung in einem weiten Sinne zur Verfügung zu stellen, entspricht ein solches Vorgehen dem Gedanken der zu Recht eingeforderten Autonomie von Studienteilnehmern in der medizinischen Forschung; das Gegenargument, mangels konkreter Informationen über künftige Studien liege keine für die Wahrung der Autonomie erforderliche „informierte" Einwilligung vor, greift zu kurz, weil es für die Betroffenen von vornherein die Möglichkeit ausschließt, bewusst ihre Daten für die gesamte medizinische Forschung zur Verfügung zu stellen.[27]

Beruht die Verarbeitung personenbezogener Daten auf einer Einwilligung, dann hat die betroffene Person nach Art. 7 Abs. 3 DSGVO das Recht, ihre Einwilligung jederzeit zu widerrufen. Das hat nach Art. 17 Abs. 1 lit. b DSGVO zur Folge, dass die betreffenden Daten zu löschen sind, sofern es an einer anderen Rechtsgrundlage für die Verarbeitung fehlt. Gerade bei KI-basierter Forschung stellt sich dann die Frage, wie eine solche Löschung vorgenommen werden kann, wenn die Daten schon in eine Auswertung eingeflossen oder etwa zum Anlernen eines neuronalen Netzes verwendet worden sind.[28] In manchen Fällen kann hier eine Anonymisierung des betreffenden zugrunde liegenden Datensatzes (z.B. durch Löschung des Eintrags der widerrufenden Person in der Pseudonymliste) das Löschen bewirken, ohne das Forschungsziel zu gefährden.[29] In anderen Fällen greift hier möglicherweise Art. 17 Abs. 3 lit. d DSGVO, wonach die Löschungspflicht nicht besteht, soweit die Erreichung des jeweiligen

26 Der Beschluss ist im Bereich „Infothek" des Internetauftritts der DSK abrufbar: https://www.datenschutzkonferenz-online.de.

27 Aus rechtsphilosophischer Sicht *T. Herbst*, Autonomie und broad consent in der medizinischen Forschung, RphZ 2019, 271 (278 ff.).

28 Dazu auch *P. Vogel*, Künstliche Intelligenz (Fn. 6), S. 91 ff.

29 Zur Anonymisierung als Form des Löschens *T. Herbst*, in: Kühling/Buchner (Hrsg.), DS-GVO/BDSG (Fn. 2), Art. 17 Rn. 39, 39a.

Forschungszwecks durch die Löschung unmöglich gemacht oder ernsthaft beeinträchtigt würde. Besteht neben der Einwilligung eine andere Rechtsgrundlage für die Verarbeitung, dann greift ohnehin schon die Löschungspflicht aus Art. 17 Abs. 1 DSGVO nicht, und die Verarbeitung kann dann auf diese andere Rechtsgrundlage gestützt werden; allerdings verlangt hier der Grundsatz der Fairness (Art. 5 Abs. 1 lit. a DSGVO, siehe dazu oben), dass der Betroffene nach Möglichkeit vor Erteilung der Einwilligung auf die Existenz dieser weiteren Rechtsgrundlage hingewiesen wurde.[30] In einigen wenigen spezialgesetzlichen Regelungen im Bereich der medizinischen Forschung finden sich besondere Rechtsgrundlagen für die Weiterverarbeitung von Daten nach einem Widerruf der Einwilligung; ein Beispiel hierfür im Arzneimittelrecht ist die Regelung in Art. 28 Abs. 3 S. 2 CTR[31] (vgl. § 40b Abs. 6 S. 2 Nr. 2 AMG). Im Übrigen kann hier (wie auch in Fällen, in denen eine Einwilligung von vornherein nicht vorhanden ist) auch an die sogenannten Forschungsklauseln in den Datenschutzgesetzen des Bundes und der Länder als Rechtsgrundlage der Verarbeitung gedacht werden; darauf soll im folgenden Abschnitt näher eingegangen werden.

c) Weitere gesetzliche Grundlagen

Fehlt es an einer Einwilligung, dann bedarf die Verarbeitung personenbezogener Daten für Zwecke der medizinischen Forschung einer anderen geeigneten gesetzlichen Grundlage. Hier zeigt sich ein Konstruktionsfehler in der föderalen Kompetenzordnung Deutschlands: Der Bund verfügt nicht über eine umfassende Gesetzgebungskompetenz etwa für die „medizinische Forschung", aufgrund derer er auch Rechtsgrundlagen für entsprechende Datenverarbeitungen schaffen könnte. Durch Bundesgesetz können nur in bestimmten einzelnen Bereichen Regelungen über die medizinische Forschung und die dafür erforderliche Datenverarbeitung geschaffen werden, so etwa im Bereich der Arzneimittel und Medizinprodukte (Art. 74 Abs. 1 Nr. 19 GG) oder in der Versorgungsforschung im Hinblick auf die Gesundheitsdaten der gesetzlich Krankenversicherten (Art. 74 Abs. 1 Nr. 12 GG). Für viele Forschungsprojekte finden sich daher mögliche Rechtsgrundlagen im Landesrecht oder in einer Kombination von Bundes- und Landesrecht,

30 Dazu näher *Herbst* (Fn. 29), Art. 17 Rn. 24a.

31 „Clinical Trial Regulation": Verordnung (EU) Nr. 536/2014 des Europäischen Parlaments und des Rates vom 16.4.2014 über klinische Prüfungen mit Humanarzneimitteln und zur Aufhebung der Richtlinie 2001/20/EG, ABl. Nr. L 158, S. 1.

wobei sich zahlreiche Unterschiede im Detail zwischen den verschiedenen landes- bzw. bundesrechtlichen Regelungen finden.[32] Im Landesrecht kommen hier etwa in Betracht die sogenannten Forschungsklauseln in den Landesdatenschutzgesetzen[33] und datenschutzrechtliche Regelungen in den Landeskrankenhausgesetzen.[34]

Diesem rechtlichen Befund steht der Umstand gegenüber, dass moderne medizinische Forschung mit großen Datenmengen in der Regel im Verbund mehrerer oder einer Vielzahl von Einrichtungen durchgeführt wird, die über mehrere Bundesländer verteilt bzw. auch im Ausland ansässig sind. Solche Verbundprojekte müssen ggf. das Datenschutzrecht mehrerer Bundesländer und möglicherweise auch bundesrechtliche Regelungen beachten. Die Notwendigkeit der gleichzeitigen Beachtung der Voraussetzungen einer Vielzahl von unterschiedlichen Rechtsgrundlagen stellt ein – in dieser Form eigentlich unnötiges – Erschwernis für solche Verbundforschung dar.

Auch die durch den Bundesgesetzgeber vor Kurzem neu geschaffene Regelung in § 287a SGB V schafft hier keine Abhilfe.[35] Nach dieser Regelung findet bei „länderübergreifenden Vorhaben der Versorgungs- und Gesundheitsforschung" § 27 BDSG Anwendung, also die Forschungsklausel im BDSG, aufgrund derer Gesundheitsdaten ohne Einwilligung der Betroffenen für Forschungszwecke verarbeitet werden dürfen, wenn das Forschungsinteresse ein Interesse der Betroffenen an einem Ausschluss der Verarbeitung erheblich überwiegt und bestimmte weitere Voraussetzungen (insbesondere die Anwendung der technischen und organisatorischen Maßnahmen des § 22 Abs. 2 BDSG) erfüllt sind. Beim ersten Hinsehen scheint diese Regelung eine Lösung des Problems zu sein: Bei länderübergreifenden Verbundprojekten scheint anstelle der unterschiedlichen landesrechtlichen Regelungen die bundesgesetzliche Forschungsklausel des § 27 BDSG einheitlich Anwendung finden zu können (wobei allerdings die Prüfung, ob das Forschungsinteresse das Interesse des Betroffenen „erheblich überwiegt", einen Unsicherheitsfaktor darstellt, weshalb solche Forschungsklauseln eher zurückhaltend angewendet werden sollten). Bei näherem

32 Dieser Befund auch bei *S. v. Kielmansegg*, Gesetzgebung im Windschatten der Pandemie: § 287a SGB V und der Datenschutz in der Gesundheitsforschung, VerwArch 2021, 133 (151 f.).

33 Z.B. in Berlin § 17 BlnDSG.

34 Z.B. in Berlin § 25 LKG.

35 Zu dieser Regelung ausführlich – allerdings im Ergebnis die Möglichkeit einer Abhilfe durch sie jedenfalls nicht ausschließend – *v. Kielmansegg*, Gesetzgebung (Fn. 32), 133 ff.

Hinsehen wird aber deutlich, dass der Anwendungsbereich des § 287a SGB V sehr beschränkt ist. Das ergibt sich zum einen aus dem Regelungsort: Das SGB V, also das Fünfte Buch des Sozialgesetzbuchs, enthält die Regelungen über die gesetzliche Krankenversicherung; schon das spricht dafür, dass § 287a SGB V nur Forschung mit den bei den gesetzlichen Krankenversicherungen vorhandenen Daten erfasst, also nicht etwa sämtliche in Kliniken vorhandenen Behandlungsdaten und auch nicht Daten der privat Versicherten. Für eine solche Beschränkung des Anwendungsbereichs des § 287a SGB V spricht im Übrigen auch die Gesetzgebungskompetenz des Bundes: Dieser hat nach Art. 74 Abs. 1 Nr. 12 GG die Gesetzgebungskompetenz für „die Sozialversicherung" und damit auch für die gesetzliche Krankenversicherung; medizinische Forschung abseits der gesetzlichen Krankenversicherung fällt aber weder unter diesen Titel noch unter den Kompetenztitel „Recht der Wirtschaft" (Art. 74 Abs. 1 Nr. 11 GG) noch unter den Titel „wirtschaftliche Sicherung der Krankenhäuser" (Art. 74 Abs. 1 Nr. 19a GG)[36] und auch nicht unter den erkennbar auf Maßnahmen wie Finanzhilfen beschränkten[37] Titel „Förderung der wissenschaftlichen Forschung" (Art. 74 Abs. 1 Nr. 13 GG).[38] Es bleibt also (leider) vorerst bei der dargestellten föderalen Rechtszersplitterung.

d) Internationale Kooperationen

KI-basierte medizinische Forschung mit großen Datenmengen kann in besonderem Maße von internationalen Kooperationen profitieren. Das betrifft nicht nur den Austausch von Fachwissen über neue Forschungsmethoden, sondern auch die Datenbasis dieser Forschung: Ein internationales Verbundprojekt kann potentiell auf einen größeren Pool von Patienten- oder Probandendaten zurückgreifen als ein auf nationale Partner beschränktes Projekt. Typischerweise werden in solchen Verbundprojekten

36 Für die beiden letztgenannten Kompetenztitel im Ergebnis ebenso *v. Kielmansegg*, Gesetzgebung (Fn. 32), 139 ff.

37 Vgl. nur *F. Wittreck*, in: H. Dreier (Hrsg.), Grundgesetz-Kommentar, Bd. III, 3. Aufl., Tübingen 2015, Art. 74 Rn. 65, wo lediglich problematisiert wird, inwieweit der Kompetenztitel auch Kontrollmaßnahmen über die Verwendung der Fördergelder umfasst.

38 Die drei letztgenannten Kompetenztitel werden in der Begründung des Gesetzentwurfs genannt: BT-Drs. 19/18111, 26. In Bezug auf die datenschutzrechtlichen Regelungen als „Förderung der wissenschaftlichen Forschung" spricht *v. Kielmansegg*, Gesetzgebung (Fn. 32), 168 von einer „unorthodoxen Auslegung", ohne diese Auslegung aber gänzlich abzulehnen.

daher zwischen den Partnern Gesundheitsdaten der Patienten bzw. Probanden ausgetauscht. Sobald dabei Partner außerhalb der EU involviert sind, entstehen besondere datenschutzrechtliche Fragen und Probleme. Die DSGVO ist nämlich so strukturiert, dass sie die Übermittlung personenbezogener Daten zwischen den Mitgliedstaaten zwar aufgrund der hier geltenden weitgehend einheitlichen datenschutzrechtlichen Anforderungen ohne besondere zusätzliche Voraussetzungen erlaubt; wegen des möglicherweise niedrigeren Datenschutzniveaus in Ländern außerhalb der EU („Drittländern") regelt die DSGVO aber in einem eigenen Kapitel (Art. 44-50 DSGVO) solche zusätzlichen Anforderungen für derartige Übermittlungen. Welche Fragen und Probleme dadurch aufgeworfen werden, soll im Folgenden anhand des für die medizinische Forschungspraxis besonders relevanten Beispiels der USA erläutert werden.

Weitgehend unproblematisch ist die Datenübermittlung, wenn für das betreffende Drittland ein sogenannter Angemessenheitsbeschluss der EU-Kommission existiert (Art. 45 DSGVO). Mit einem solchen Angemessenheitsbeschluss bescheinigt die Kommission dem betreffenden Drittland gewissermaßen, dass in diesem Land ein der EU vergleichbares Datenschutzniveau besteht. Für die USA existierten in der Vergangenheit zwei solche Angemessenheitsbeschlüsse („Safe Harbor" und „Privacy Shield"), die allerdings beide durch den EuGH (in den Entscheidungen „Schrems I" und „Schrems II") für nichtig erklärt wurden.[39] Der EuGH befand das Datenschutzniveau in den USA vor allem deswegen als nicht ausreichend, weil nach amerikanischem Recht die Möglichkeit besteht, dass die Geheimdienste auf personenbezogene Daten von Ausländern zugreifen, ohne dass die betroffenen Personen ausreichende Rechtsschutzmöglichkeiten dagegen haben. Nach erneuten Verhandlungen mit den USA hat jüngst die Kommission den dritten Angemessenheitsbeschluss erlassen („EU-US Data Privacy Framework").[40] Ob nun dieser Angemessenheitsbeschluss den Anforderungen des EuGH standhält, ist allerdings nicht gewiss, so dass nicht ausgeschlossen ist, dass auch er durch den EuGH für nichtig erklärt werden wird. Ein Angemessenheitsbeschluss ist daher nach den bisherigen Erfahrungen jedenfalls gegenwärtig nicht als dauerhaft sichere Rechtsgrundlage für Übermittlungen in die USA zu betrachten.

39 Safe Harbor: EuGH DVBl 2015, 1446 („Schrems I"); dazu *C. Schröder*, in: Kühling/Buchner (Hrsg.), DS-GVO/BDSG (Fn. 2), Art. 45 Rn. 41 ff. Privacy Shield: EuGH DVBl 2020, 1259 („Schrems II"); dazu *Schröder*, ebda., Art. 45 Rn. 44 ff.

40 Angemessenheitsbeschluss der Kommission vom 10.7.2023, Dokument Nr. C(2023) 4745 final.

Fehlt es an einem Angemessenheitsbeschluss, dann bietet Art. 46 DSGVO weitere Instrumente für Übermittlungen in Drittländer. Hier kommen insbesondere die Standarddatenschutzklauseln (auch Standardvertragsklauseln genannt) gemäß Art. 46 Abs. 2 lit. c DSGVO in Betracht. Hierbei handelt es sich um von der Kommission beschlossene Textbausteine, die in einem Kooperationsvertrag mit einem Datenempfänger im Drittland verwendet werden können. Gerade im Hinblick auf die USA hat der EuGH in der Entscheidung „Schrems II" allerdings festgestellt, dass die Verwendung dieser Standarddatenschutzklauseln für sich genommen noch nicht ein ausreichendes Datenschutzniveau garantiert; das liegt vor allem daran, dass durch einen Vertrag zwischen zwei Kooperationspartnern die Zugriffsmöglichkeiten staatlicher Behörden auf Daten nicht wirksam beschränkt werden können. Der EuGH fordert daher bei der Verwendung von Standarddatenschutzklauseln zusätzliche Sicherungsmaßnahmen, wobei aber nicht verbindlich geklärt ist, welche zusätzlichen Sicherungsmaßnahmen letztlich ausreichend sind. Als zusätzliche Maßnahmen im Bereich der KI-basierten medizinischen Forschung ist hier etwa zu denken an die Anonymisierung der Daten, Verwendung synthetischer Daten (also Daten, die zwar – z.B. mit Hilfe von sog. „Generative Adversarial Networks" – künstlich erzeugt sind, aber bei der Auswertung zum selben Ergebnis wie die Originaldaten führen)[41] oder an Techniken wie „differential privacy" (Veränderung der Ausgangsdaten unter Beibehaltung ihrer statistischen Aussagekraft),[42] homomorphe Verschlüsselung (Verschlüsselung der Ausgangsdaten in einer Weise, die eine Auswertung der verschlüsselten Daten ohne Entschlüsselung ermöglicht) oder „code to data" (die Übermittlung von Auswertungssoftware an den Ort, an dem die Daten vorhanden sind, anstelle der Übermittlung der Daten).[43]

Schließlich gibt es auch die in Art. 49 Abs. 1 lit. a DSGVO vorgesehene Möglichkeit, dass die betroffene Person in die Datenübermittlung in das Drittland ausdrücklich einwilligt, nachdem sie über die möglichen Risiken unterrichtet wurde. Hier ist zu beachten, dass die betreffenden Risiken nicht bloß pauschal angesprochen werden, sondern möglichst konkret, so dass die betroffene Person sich eine Vorstellung davon machen kann, was

41 Dazu *R. Behrang*, Rechtliche Bewertung synthetischer Daten für KI-Systeme, DuD 2021, 303 (305).

42 Dazu *F. Boenisch*, Privatsphäre und Maschinelles Lernen, DuD 2021, 448 (450 ff.).

43 Vgl. zu solchen Techniken auch *C. Winter/V. Battis/O. Halvani*, Herausforderungen für die Anonymisierung von Daten, ZD 2019, 489 (490 ff.); *Vogel*, Künstliche Intelligenz (Fn. 6), S. 221 ff.

es bedeutet, dass ihre Daten in das betreffende Drittland übertragen werden. Es müsste also etwa darüber informiert werden, dass die Möglichkeit besteht, dass staatliche Behörden auf die Daten zugreifen, ohne dass die betroffene Person ausreichende Rechtsschutzmöglichkeiten dagegen hat.

IV. Fazit

Die obigen Ausführungen haben gezeigt, dass KI-basierte medizinische Forschung mit den Grundsätzen der DSGVO im Wesentlichen vereinbar ist. Die wenigsten Probleme bereiten hierbei Vorgehensweisen, bei denen die Patienten bzw. Probanden in die Verarbeitung ihrer Daten zu Forschungszwecken einwilligen.

Fehlt es an solchen Einwilligungen (was z.B. der Fall sein kann, wenn die betreffenden Daten zunächst in einem Behandlungskontext angefallen sind und erst später ihre Forschungsrelevanz offenbar wird), dann bedarf die Verarbeitung der betreffenden medizinischen Daten einer anderen Rechtsgrundlage. Insbesondere für Verbundprojekte wäre hier jedenfalls für Deutschland eine einheitliche Rechtsgrundlage (anstelle der bestehenden föderalen Vielfalt) wünschenswert. Allerdings fehlt dem Bund hierfür, wie gezeigt, eine umfassende Gesetzgebungskompetenz etwa für die medizinische Forschung. Abhilfe könnte insoweit durch eine Änderung des Grundgesetzes geschaffen werden, die allerdings wegen der erforderlichen verfassungsändernden Mehrheiten in Bundestag und Bundesrat in politischer Hinsicht eine hohe Hürde darstellt. Möglicherweise lässt sich eine solche Änderung in eine umfassendere Neujustierung des Föderalismus einbetten. Eine rechtliche Alternative zur Änderung des Grundgesetzes wäre etwa ein Staatsvertrag, den sämtliche Bundesländer und der Bund abschließen und in dem sich Bund und Länder auf eine einheitliche Regelung zur datengestützten Forschung einigen. Allerdings stellt auch der Abschluss eines solchen Staatsvertrages eine hohe politische Hürde dar.