Regina T. Riphahn, Andreas Peichl, Bertram Raum, Kerstin Schneider und Daniel Vorgrimler

Ein Forschungsdatengesetz für Deutschland: Datenzugang, Datenverknüpfung und Datenschutz

Übersicht

I. Einleitung

II. Was muss im Forschungsdatengesetz kurzfristig geregelt werden?

1. Kernforderungen

a) Datenzugang verbessern

b) Datenverknüpfung ermöglichen

c) Koordination eines ermöglichenden Datenschutzes

2. Forschungsgeheimnis, Beschlagnahmeverbot und Zeugnisverweigerungsrecht

3. Reform des Bundesstatistikgesetzes

4. Einrichtung eines Treuhänders für Mikrodaten

III. Deutschland in europäischen Forschungsdatenräumen

IV. Weitere drängende Themen

1. Bildungsverlaufsregister

2. Prozessdaten von Unternehmen nutzbar machen

V. Fazit

Deutschland braucht ein Forschungsdatengesetz. Diese Forderung wurde in den Koalitionsverträgen der 20. und 21. Legislaturperiode aufgegriffen, da die rechtliche Ausgestaltung der Forschungsdateninfrastruktur in Deutschland weder den Anforderungen der internationalen Forschung noch denen einer modernen evidenzbasierten Politik und Verwaltung entspricht. Forschungsdaten fehlen auch zur Bewältigung gesellschaftlicher Herausforderungen in akuten Krisen. Der Bedarf an besseren Forschungsdaten ist unbestritten. Es hängt jedoch von der konkreten Ausgestaltung eines Forschungsdatengesetzes ab, ob die bestehenden Defizite abgebaut werden können.

Dieser Beitrag erläutert zunächst drei Kernforderungen, die durch ein Forschungsdatengesetz unmittelbar adressiert werden müssen. Dies sind die Verbesserung des Datenzugangs, das Zusammenspielen von Daten und der Kulturwandel hin zu einem ermöglichenden Datenschutz. Neben diese Kernforderungen tritt das vom Bundesverfassungsgericht unterstützte Erfordernis der Einrichtung eines Forschungsgeheimnisses, die auch unabhängig vom Forschungsdatengesetz geforderte Novellierung und Modernisierung des Bundesstatistikgesetzes sowie die Einrichtung eines Datentreuhänders.

Aber nicht nur der nationale Gesetzgeber gestaltet die Verfügbarkeit von Forschungsdaten. Auf Ebene der Europäischen Union werden Konzepte für einen europäischen Forschungsraum und eine europäische Forschungsdateninfrastruktur erarbeitet. Hier ist die Bundesregierung aufgerufen, die Weichenstellung in Europa konstruktiv mitzugestalten.

Es gibt zudem Desiderate, die über die Kernforderungen an ein Forschungsdatengesetz hinausgehen, aber Teil der Weiterentwicklung der Forschungsdateninfrastruktur sein müssen. Dazu gehören der Aufbau eines Bildungsverlaufsregisters sowie die Unterstützung von Projekten zum Datenaustausch zwischen Wirtschaft und Wissenschaft.

I. Einleitung

Um Entwicklungen in Wirtschaft und Gesellschaft zu analysieren und auf dieser Basis Politik und Öffentlichkeit zu informieren und zu beraten, benötigt die Wissenschaft gute Daten.[1] Diese sind in Deutschland oft nicht nutzbar. Deutschland braucht ein Forschungsdatengesetz, um die bestehenden Defizite abzubauen. Diese Forderung wurde in den Koalitionsverträgen der 20. und 21. Legislaturperiode aufgegriffen, denn die rechtliche Ausgestaltung der Forschungsdateninfrastruktur in Deutschland genügt aktuell weder den Anforderungen der internationalen Forschung noch denen einer modernen evidenzbasierten Politik und Verwaltung.[2] Die Notwendigkeit besserer Forschungsdaten ist unstrittig. Es kommt nun aber auf die konkrete Ausgestaltung des Gesetzes an, wenn die bestehenden Defizite abgebaut werden sollen.

1 Die hier relevanten Forschungsdaten sind potentiell sensible, z.T. personenbezogene (Einzel-/Individual-)Daten, die von verschiedenen Einrichtungen für unterschiedliche Zwecke erzeugt werden.

2 Vgl. *Riphahn*, Datenzugang für die Forschung – Eine Einleitung, Wirtschaftsdienst 103(11), 728.

Die fehlenden Forschungsdaten haben schwerwiegende Konsequenzen für Politik, Wissenschaft und Verwaltung. Erstens müssen in Deutschland immer wieder und gerade auch in Krisensituationen uninformiert und ohne empirische Evidenz Entscheidungen getroffen werden. So wurde in der Corona-Pandemie Evidenz aus Israel und Großbritannien herangezogen, obwohl die Übertragbarkeit internationaler Studien auf Deutschland fraglich ist. Mangels Daten mussten wirtschaftspolitische Beschlüsse wie im Blindflug getroffen werden: die Unterstützung von Unternehmen während der Corona-Pandemie erfolgte unpräzise. Auch in der Gaspreiskrise konnte nicht zielgenau agiert werden. Aber nicht nur in Krisensituationen erschweren fehlende Daten fundierte Entscheidungen. Die Simulationsrechnungen zur Vorbereitung der Kindergrundsicherung im Jahr 2024 erfolgten auf Basis von Daten aus dem Jahr 2018. Dabei ist klar: Politik ist nur dann zielgenau und effizient, wenn sie sich auf aktuelle und verlässliche wissenschaftliche Evidenz stützen kann.[3]

Zweitens ist der schlechte Datenzugang ein erheblicher Nachteil für den Wissenschaftsstandort Deutschland. Eine Befragung des Vereins für Socialpolitik, der größten wirtschaftswissenschaftlichen Fachgesellschaft im deutschsprachigen Raum, ergab 2023, dass 65% der antwortenden Wissenschaftler:innen unzufrieden mit dem Zugang zu Forschungsdaten in Deutschland sind. 73 % der Befragten geben an, dass ihr Forschungserfolg unter den Restriktionen beim Datenzugang leidet. Etwa zwei Drittel der Antwortenden konnte in den letzten 10 Jahren Projekte aufgrund von Auflagen und Restriktionen beim Datenzugang nicht im gewünschten Umfang durchführen. 75% der Befragten sind der Ansicht, dass der Datenzugang für die Forschung in Deutschland eine fundierte und aktuelle Politikberatung erschwert. In Relation zu vergleichbaren Ländern wird der Datenzugang für die Forschung in Deutschland von 78,5% der Befragten als schlechter eingeschätzt.[4]

Diese im internationalen Vergleich schlechteren Forschungsbedingungen führen dazu, dass deutsche Forschende mit Daten aus anderen Ländern arbeiten müssen. Dadurch fördern die von deutschen Steuerzahlenden finanzierten Forschungseinrichtungen die Erforschung und Beratung unserer Nachbarländer, während in Deutschland zu relevanten, vergleichbaren Problemlagen wenig bekannt ist. Zudem ist Deutschland als Forschungsstandort ohne wettbewerbsfähigen Datenzugang für die besten Forschenden nicht attraktiv: Sie gehen dorthin, wo sie am besten forschen können. Und das ist eindeutig nicht in Deutschland.

Drittens kann auch die Verwaltung in Deutschland ihre Aufgaben nicht effizient erfüllen. Unter anderem, um den EU-Datenanforderungen entsprechen zu können, plant Deutschland im Jahr 2031 einen Registerzensus durchzuführen, der ohne breite Befragungen auskommt. Ob das föderale deutsche Einwohnermeldewesen dazu in der Lage sein wird, wird weithin bezweifelt.

Eine vierte Konsequenz mangelnder Daten betrifft die Fähigkeit des Staates, die Wirksamkeit der eigenen Maßnahmen bewerten zu können. Die Möglichkeiten, ex ante und ex post Evaluationen von Gesetzen durchzuführen, sind in Deutschland besonders aufgrund der Restriktionen bei Datenzugang und Datennutzung eng begrenzt. Wenn Politik und Verwaltung aber nicht aus Erfahrung lernen können, bleiben Chancen ungenutzt, Maßnahmen effektiver und effizienter zu gestalten.

Die Forschungsdatenstruktur in Deutschland ist im Vergleich zu unseren Nachbarländern (u.a. Österreich, Schweiz, Frankreich, Niederlande, Dänemark) rückständig. Es fehlt an vielem, was für evidenzbasierte Politik und moderne Wissenschaft unverzichtbar ist, von der Datenerhebung, über die Datenbereitstellung, bis zur Datenverknüpfbarkeit. Um in Europa aufschließen zu können und die Hindernisse für eine Weiterentwicklung der Dateninfrastruktur abzubauen, müssen in Deutschland dringend die gesetzlichen Grundlagen angepasst werden.

Insbesondere die fehlende Datenverknüpfbarkeit ist ein Problem im dezentral organisierten deutschen Datenökosystem. Bildhaft kann man sich die deutsche Forschungsdatenlandschaft als unverbundene Datensilos vorstellen. Manche dieser Silos, wie das für Arbeitsmarktdaten, sind sehr gut gefüllt, andere, wie das für Steuerdaten wurden kürzlich gestärkt[5], wiederum ande-

3 Vgl. *Riphahn, Schneider,* „Für einen besseren Zugang zu Forschungsdaten", FAZ vom 11.12.2023.

4 Vgl. *Chaskel, Getzner, Fürnkranz-Prskawetz, Riphahn, Schmidheiny,* Zugang zu Forschungsdaten in den D-A-CH-Ländern: Eine Vermessung der (Un-)Zufriedenheit, Perspektiven der Wirtschaftspolitik 25(1) 2024, 5-19.

5 In Folge des Gutachtens des Wissenschaftlichen Beirats beim BMF „Notwendigkeit, Potenzial und Ansatzpunkte einer Verbesserung der Dateninfrastruktur für die Steuerpolitik" (2020) wurde u.a. das „Netzwerk empirische Steuerforschung (NeSt)" als Plattform zur Vernetzung von Wissenschaftlerinnen und Wissenschaftlern im Bereich der empirischen Steuerforschung mit der amtlichen Statistik und der Finanzverwaltung gegründet. Zudem wurden Ressourcen beim statistischen Bundesamt ausgeweitet und das Angebot an Steuerdaten vergrößert.

re, wie das Silo mit Gesundheitsdaten, werden hoffentlich bald befüllt. Im Bereich der Gesundheitsdaten wurde im Jahr 2024 im Vorgriff auf den Europäischen Gesundheitsdatenraum durch das Gesundheitsdatennutzungsgesetz (GDNG) vom 25.03.2024[6] ein (kleiner) Fortschritt erzielt: Gesundheitsdaten aus dem Forschungsdatenzentrum (FDZ) Gesundheit beim BfArM dürfen nun mit anderen Gesundheitsdaten (hier: Krebsregisterdaten) verknüpft werden.

Andere, wie das Bildungsdatensilo, sind noch erschreckend leer. Dass Daten bereichsspezifisch aufbereitet und gehalten werden, ist sinnvoll. Die Daten müssen aber auch, wie es international üblich ist, über Bereiche hinweg verknüpft werden können. Dem stehen in Deutschland derzeit noch zahlreiche gesetzliche Regelungen entgegen.

Es ist nicht einfach, eine über Jahrzehnte gewachsene, aber nun veraltete Dateninfrastruktur und ihre rechtliche Grundlage zu modernisieren. Will man nicht noch weiter zurückfallen, besteht dringender Handlungsbedarf. Dazu müssen zwei grundsätzliche Probleme gelöst werden. Zum einen haben wir in Deutschland keine Identifikationsnummern, um Daten aus verschiedenen Quellen zusammenzuspielen. In unseren Nachbarländern können Daten aus verschiedenen Silos hingegen verknüpft werden. Zum anderen führen in Deutschland zahlreiche Einzelgesetze zur Nutzung von Daten dazu, dass jede Datennutzung verboten ist, die nicht explizit erlaubt ist. Ausnahmen für Wissenschaft und Forschung gibt es nur selten – und wenn dann für eng begrenzte Spezialfälle. Dabei würde das in der Europäischen Datenschutzgrundverordnung (DSGVO) festgeschriebene Forschungsprivileg auch in Deutschland bereits heute deutlich mehr ermöglichen.

Um in diesem Themenbereich voran zu kommen, sah der Koalitionsvertrag der „Ampelregierung" von 2021 ein Forschungsdatengesetz (FDG) vor, das die Datengrundlage für Wissenschaft und Forschung verbessern sollte. Auch der Koalitionsvertrag von Union und SPD verspricht ein Forschungsdatengesetz noch im Jahr 2025. Im europäischen Ausland gibt es viele Vorbilder für einen wissenschaftsadäquaten Zugang zu Forschungsdaten. Entscheidend ist, dass das FDG drei Kernprobleme löst: 1. Datenzugang für die Forschung schaffen, 2. Datenverknüpfung erlauben und 3. einen forschungsermöglichenden Datenschutz etablieren.

In diesem Beitrag sollen die Kernprobleme und Lösungsansätze erörtert werden. Zum ersten Punkt gehört, dass der Zugang zu öffentlichen Verwaltungs- und Registerdaten für die Forschung prinzipiell ermöglicht werden muss und zeitgemäße, sichere Remote-Desktop-Datenzugangsmöglichkeiten für die Wissenschaft umgesetzt werden müssen. Für den zweiten Punkt, die Datenverknüpfung, müssen Daten der öffentlichen Hand, Daten aus Forschungsinstituten und der Wirtschaft für Zwecke der unabhängigen Forschung unter Verwendung einer Identifikationsnummer datenschutzkonform zusammengespielt werden. Dies kann – wie im europäischen Ausland üblich – über einen Datentreuhänder erfolgen. Um den dritten Punkt anzugehen, muss die zersplitterte Anwendung von Datenschutzregeln im föderalen Deutschland hinterfragt werden. Projekte sollten nicht von 16+1 Datenschutzaufsichtsbehörden geprüft werden müssen. Stattdessen sollten Regeln bundeseinheitlich ausgelegt werden und Forschung ermöglichen.

Alle drei Forderungen betreffen auch die amtliche Statistik. Die gesetzliche Grundlage der amtlichen Statistik, das aktuelle Bundesstatistikgesetz (BStatG), atmet den Geist der 1980er Jahre und sollte novelliert werden.[7] Eine moderne amtliche Statistik muss den Informationsbedarf von Gesellschaft und Politik effektiv decken können. Statt der amtlichen Statistik einzelne Erhebungsfragen gesetzlich vorzuschreiben (Inputorientierung), ist auch als Beitrag zum Bürokratiekostenabbau eine zeitgemäße amtliche Statistik über ihre Arbeitsergebnisse (Outputorientierung) zu regulieren. Dabei ist die Kooperation zwischen amtlicher Statistik und Wissenschaft zu nutzen und zu fördern.

Die Aufnahme des Forschungsdatengesetzes (FDG) in die Koalitionsverträge der 20. und 21. Legislaturperiode zeigt, dass Handlungsbedarf gesehen wurde. Das Bundesministerium für Bildung und Forschung wurde 2021 beauftragt, einen Gesetzentwurf für das FDG zu erarbeiten. Der Prozess wurde von der Wissenschaft inten-

6 Gesetz zur verbesserten Nutzung von Gesundheitsdaten (GDNG) v. 25. 3. 2024, BGBl. 2024 I Nr. 102.

7 Das Bundesstatistikgesetz (BStatG) von 1953 wurde nach dem Volkszählungsurteil vom 15.12.1983 ebenso wie die Fachstatistikgesetze novelliert. Das derzeitige BStatG vom 20.10.2016 (BGBl. I S. 2394) basiert auf dem BStatG vom 22.1.1987 (BGBl. I S. 462).

siv unterstützt und begleitet.[8] Der im Herbst 2024 vorgelegte Gesetzentwurf wurde nicht mit allen Akteuren abgestimmt, bietet aber eine Grundlage, um den Zugang zu Forschungsdaten für die Wissenschaft zu verbessern und das Vorhaben zeitnah abzuschließen.

Der gesellschaftliche Mehrwert der Modernisierung des Forschungsdatenzugangs mag zunächst abstrakt und nicht unmittelbar sichtbar und messbar sein, er ist aber immens. Es geht zum einen darum, transparente und informierte politische Entscheidungen im nationalen Interesse zu ermöglichen und durch datengestütztes Handeln Resilienz in Krisensituationen zu gewährleisten. Die Kosten fehlender Daten und ungenutzter wissenschaftlicher Potenziale sind hoch. Eine kleine, evidenzbasierte Steigerung der Effizienz von Förderprogrammen spart ein Vielfaches der für die Datenbereitstellung benötigten Ressourcen.[9] Es geht aber auch darum, hochklassige und international wettbewerbsfähige Forschung in Deutschland zu ermöglichen.

Deutschland muss hier keine neuen Wege gehen, sondern kann sich an international üblichen, rechtssicheren und datenschutzkonformen Ansätzen der Datennutzung und -verknüpfung orientieren. Die Vorbilder beispielsweise in Frankreich, Österreich, und den Niederlanden zeigen, dass sich Datenschutz mit zeitgemäßer Datennutzung im Gemeinwohlinteresse vereinen lassen.

II. Was muss im Forschungsdatengesetz (FDG) geregelt werden?

1. Kernforderungen

Forschende monieren zahlreiche Defizite im deutschen Forschungsdatenökosystem. Ein FDG kann hier Lösungen anbieten. Im Folgenden werden die drei zentralen Aspekte diskutiert, die ein FDG regeln muss:[10] Der Datenzugang muss verbessert werden, Daten müssen verknüpft werden und Datenschutz muss Forschung ermöglichen.

a) Datenzugang verbessern

Im Gegensatz zu anderen Ländern fehlt in Deutschland für die Forschung die grundsätzliche Möglichkeit eines Zugangs zu öffentlichen Verwaltungs- und Registermikrodaten. Mikrodaten bieten Einzelfallinformationen zu Personen, Haushalten, Betrieben oder Unternehmen. Viele dieser Daten sind zwar vorhanden, aber für die Forschung nicht oder nur sehr eingeschränkt zugänglich.

Ein Beispiel für Daten aus Verwaltungsprozessen ist das Unternehmensbasisdatenregistergesetz (UBRegG), das die Forschung bisher nicht berücksichtigt. Lediglich für Statistikzwecke (§ 281 SGB III) und für Zwecke der Betriebsdatenführung (§ 18i Abs. 6 SGB IV) können die Informationen durch die Bundesagentur für Arbeit verwendet werden. Damit ist die Nutzung des Basisregisters stark eingeschränkt und für die Forschung nicht zugänglich.

Neben den rechtlichen Möglichkeiten fehlt häufig auch die grundsätzliche Bereitschaft öffentlicher Stellen, existierende Daten zu Forschungszwecken bereitzustellen. Im FDG sollte der Gesetzgeber öffentliche Stellen verpflichten, vorliegende amtliche Daten für wissenschaftliche Zwecke auf Mikroebene verfügbar zu machen.

Dabei ist es sinnvoll, eine generelle Lösung zu finden, die ein grundsätzliches Forschungsprivileg beim Datenzugang etabliert und möglichst nicht den unübersichtlichen Weg über Sonderregelungen in Einzelgesetzen geht. Im Gegensatz zur heutigen restriktiven Regelung sollte erlaubt sein, was nicht verboten ist.

8 Vgl. RatSWD (2022). Eckpunkte für ein Forschungsdatengesetz, abrufbar unter https://www.konsortswd.de/wp-content/uploads/RatSWD-Positionspapier-Eckpunkte-fuer-ein-Forschungsdatengesetz.pdf; *Riphahn*, Wir wissen in Deutschland vieles nicht, was wir wissen sollten, 2022, Perspektiven der Wirtschaftspolitik 23(1), 38-48; RatSWD (2023). Forschungsdatengesetz: Was zentral ist, abrufbar unter https://www.konsortswd.de/wp-content/uploads/RatSWD_Positionspapier_Forschungsdatengesetz_Was_zentral_ist.pdf, *Riphahn*, (2023). Die deutsche Dateninfrastruktur aus Sicht der empirischen Wirtschaftsforschung, Wirtschaftsdienst 103(1), 24-28; *Riphahn*, (2023). Datenzugang für die Forschung – Eine Einleitung; Zeitgespräch: Datenzugang für Forschung und Politikberatung in Deutschland, Wirtschaftsdienst 103(11), 728; RatSWD (2024), Stellungnahme zu den BMBF-Eckpunkten zum Forschungsdatengesetz, abrufbar unter https://www.konsortswd.de/wp-content/uploads/RatSWD-Stellungnahme-Eckpunkte-Forschungsdatengesetz.pdf; *Riphahn, Schneider*, Ein Forschungsdatengesetz jetzt!, 2024, Wirtschaftsdienst 104(10), 658-659; RatSWD (2025). Datensilos aufbrechen: Forschung braucht gute Daten! Forderungen des RatSWD für die 21. Legislaturperiode; abrufbar unter https://www.konsortswd.de/wp-content/uploads/RatSWD-Forderungen-21-LP.pdf (alle zuletzt abgerufen am 08.04.2025).

9 Für ein Zahlenbeispiel vgl. Sachverständigenrat zur Beurteilung der gesamtwirtschaftlichen Entwicklung, Jahresgutachten 2023/24, Kapitel 6, Textziffer 554.

10 Eine ausführliche Darstellung findet sich in *Gathmann, von Gaudecker, Riphahn*, Arbeitsmarktdaten: Gesetzesänderungen gefordert, Zeitgespräch: Datenzugang für Forschung und Politikberatung in Deutschland, 2023, Wirtschaftsdienst 103(11), 729-732; *Altmann et al.*, Für einen besseren Datenzugang – Stellungnahme der Arbeitsgruppe Arbeitsmarkt- und Sozialversicherungsdaten. Verein für Socialpolitik Berlin, abrufbar unter https://www.socialpolitik.de/sites/default/files/2023-11/1_Stellungnahme_UAG_Arbeitsmarktdaten_0.pdf (zuletzt abgerufen am 08.04.2025).

b) Datenverknüpfung ermöglichen

Viele Forschungsfragen sind auf Verwaltungsdaten aus unterschiedlichen Quellen angewiesen.[11] Ein Beispiel für solche Verknüpfungsbedarfe ist die Forschung zu den Auswirkungen der Mindestlohngesetzgebung. Hier bieten die Registerdaten der Bundesagentur für Arbeit eine wichtige Basis; allerdings enthalten diese weder Informationen über die Anzahl der individuellen Arbeitsstunden noch über die personelle Zusammensetzung von Haushalten. Daher kann beispielsweise nicht überprüft werden, ob Mindestlohnanpassungen Armut reduziert. Auch der Effekt von Mindestlohnanpassung auf die Gesundheit der unteren Einkommensgruppen kann ohne eine Verknüpfung mit Gesundheitsdaten nicht analysiert werden.

Bisher ist in Deutschland eine umfassende Verknüpfung getrennter Datenerhebungen regelmäßig nicht erlaubt. Stattdessen werden separate Erhebungen zu spezifischen Fragestellungen durchgeführt, obwohl die europäischen Rahmengesetze aufgrund des Forschungsdatenprivilegs (Art. 89 und 5 Abs. 1 lit. b DSGVO) und der Verarbeitungsgrundlagen (Art. 6 Abs. 1 lit. a-f DSGVO) die Zusammenführung der Daten grundsätzlich zulassen.

Die Zusammenführung existierender Daten aus verschiedenen Quellen bietet neben der Beantwortung gesellschaftlich relevanter Fragen weitere Vorteile: Die Verlässlichkeit der Informationen und die Datenqualität werden verbessert, auf die mehrfache Erhebung gleicher Informationen bei gleichen Erhebungseinheiten kann verzichtet werden, was die Antwortlast (response burden) für die Befragten verringert und das once-only Prinzip umsetzt.

Neben der Frage nach der rechtlichen Zulässigkeit der Datenverknüpfung stellt sich eine methodische Herausforderung, da für die korrekte Verknüpfung auf der Mikroebene von Personen, Haushalten, Betrieben oder Unternehmen Kennnummern (Identifikatoren) in den entsprechenden Datensätzen benötigt werden. In den Nachbarländern Deutschlands werden für die Verknüpfung statistische Identifikatoren (z.B. die Steueridentifikationsnummer einer Person) genutzt. Österreich sieht nach § 2d Forschungsorganisationsgesetz für die Forschung die Möglichkeit eines bereichsspezifischen Personenkennzeichens „Forschung" (bPK-BF-FO) vor. Alternativ kann ein persönlicher Identifikator genutzt werden, der bei Geburt oder Zuwanderung zugewiesen wird und für verschiedene Verwaltungszwecke hilfreich sein kann. Die rechtliche Zulässigkeit solcher Identifikatoren in Deutschland ist noch offen.[12]

Die grundsätzliche Diskussion über die rechtliche Zulässigkeit von Datenverknüpfungen hängt an der Auslegung und Umsetzung von Datenschutzvorgaben. Nach höchstrichterlicher Rechtsprechung soll vermieden werden, dass eine datenhaltende Stelle aus der Kombination von Daten aus verschiedenen Quellen Profile von Individuen, Haushalten oder Unternehmen erstellen kann.[13] Im europäischen Ausland wird dieses Problem durch die Institution des Datentreuhänders gelöst. Hinter sicheren datenschutzrechtlichen „Schutzwällen" darf nur eine regulierte Treuhandstelle Daten mit Identifikatoren aus verschiedenen Quellen zu Forschungszwecken zusammenführen. Nach der Zusammenführung werden die Identifikatoren gelöscht und die Daten der Forschung faktisch anonymisiert und unter strengen datenschutzrechtlichen Vorgaben zur Verfügung gestellt. Einen solchen pragmatischen und zielführenden Weg sah auch der Gesetzentwurf zum FDG aus dem Jahr 2024 für Deutschland vor.

In einem FDG muss Forschung als Zweck von Datenverknüpfung rechtskreisübergreifend gesetzlich verankert werden, um den Zugang zu und die Verknüpfung von Daten verschiedener Anbieter zu Forschungszwecken zu ermöglichen.

c) Koordination eines ermöglichenden Datenschutzes

Die im internationalen Vergleich strikte Interpretation von Datenschutzregeln in Deutschland und ihre unterschiedliche föderale Anwendung sind wesentliche Hindernisse für wissenschaftliches Arbeiten.

Der Datenschutz ist neben der DSGVO durch unterschiedliche nationale gesetzliche Vorgaben geregelt. Diese Regelungen sind nicht primär nach Sachgründen, wie der Art der Erhebung oder der Materie, sondern nach Rechtskreisen in zahlreichen Spezialgesetzen geregelt. Bundesländerübergreifende Vorhaben werden derzeit durch oftmals voneinander abweichende Landesdaten-

11 Dieser Umstand ist auch für die Bewertung der Regelungen im GDNG relevant, das Verknüpfungen mit nicht-medizinischen Daten nicht berücksichtigt.

12 *Martini, Wagner, Wenzel* (2017). Rechtliche Grenzen einer Personen- bzw. Unternehmenskennziffer in staatlichen Registern, Gutachten für den Nationalen Normenkontrollrat. Vgl. https://www.normenkontrollrat.bund.de/Webs/NKR/SharedDocs/Downloads/DE/Gutachten/2017-erg%C3%A4nzendes-gutachten-speyer-beistellung-pkz.pdf?__blob=publicationFile&v=2.

13 BVerfGE 65, 1 (53) unter Berufung auf BVerfGE 27, 1 (6).

schutzregelungen oder -auslegungen beeinflusst. Derzeit müssen in der Regel[14] Datenschutzbeauftragte aller Bundesländer einzeln zustimmen, wenn spezifische Informationen für die Durchführung eines Forschungsprojekts erforderlich sind. Ein Beitrag des FDG sollte darin bestehen, bundeseinheitliche und rechtskreisübergreifende Datenschutzregeln für die Datennutzung zu Forschungszwecken zu etablieren.

Um den Datenschutz zu gewährleisten, ist die Datenbereitstellung zu Forschungszwecken in Deutschland zudem oft an inhaltlich definierte Forschungsthemen gebunden. Hierdurch entstehen eng verstandene Zweckbindungen, welche die Datennutzung auf Aufgabenbereiche einzelner Verwaltungen beschränken und ihre Auswertung für Forschungsvorhaben und Politikberatung stark begrenzt oder sogar unmöglich macht. Solche Zweckbindungen für Forschungsdaten sollten gestrichen werden.

Ähnliches gilt für gesetzliche Löschfristen: Derzeit sehen zahlreiche gesetzliche Regelungen die Löschung von Informationen nach Ablauf einer Frist als Instrument des Datenschutzes vor. Zum einen ist es wenig intuitiv, dass gerade veraltete Informationen durch Löschungen geschützt werden sollen. Zum anderen liegen die hohen Kosten einer Vernichtung von Information auf der Hand. Wirtschafts- und sozialhistorische Forschung wird damit unmöglich gemacht. Ein Lernen aus alten Krisen und Reformen wird systematisch verhindert.

Gerade wenn identifizierende Informationen gelöscht werden müssen, verhindert dies endgültig eine Verknüpfung mit anderen Datensätzen oder die Nutzung von Längsschnittinformationen bei Wiederholungsbefragungen, selbst wenn sich politische oder rechtliche Rahmenbedingungen und Zielsetzungen verändern. Das Erkenntnispotenzial von Erhebungen wird dadurch erheblich gemindert. Es ist offensichtlich, dass Verknüpfungsidentifikatoren oder Hilfsmerkmale (z.B. des Registerzensus), die für Erzeugung und Verknüpfung von Forschungsdaten erforderlich sein können, gesetzlich vor Löschung geschützt werden sollten. Sie können getrennt gespeichert werden, so dass etwaigem Missbrauch effektiv vorgebeugt wird.

Wünschenswert ist auch eine einheitliche Regelung des Datenzugangs im Bereich der amtlichen Statistik. Nach § 16 Abs. 6 Nr. 2 BStatG ist der Zugriff auf formal anonymisiertes Datenmaterial im Rahmen von Gastaufenthalten in einem FDZ der Statistischen Ämter des Bundes (FDZ destatis) oder der Länder möglich. Dies betrifft aber nicht alle amtlichen Mikrodaten. Ein Beispiel sind Wirtschaftsstatistiken. Hier ist der Zugang zu formal anonymisiertem Datenmaterial an Gastwissenschaftsarbeitsplätzen nicht für alle Bundesländer gleichermaßen gegeben, da die Normen in den statistischen Landesämtern unterschiedlich interpretiert werden. Durch die uneinheitliche Datenbereitstellung entsteht vermeidbarer Mehraufwand bei den Nutzenden und bei den Statistischen Ämtern.

Zur Entlastung der Statistischen Landesämter könnte ebenfalls eine zentrale Stelle vorgesehen werden, die die datenschutzrechtliche Prüfung von Forschungsprojekten übernimmt.[15] Hinsichtlich der Datenschutzaufsicht bei länderübergreifenden Forschungsvorhaben findet sich in § 5 GDNG eine richtungsweisende Regelung. Danach ist eine federführende Datenschutzaufsichtsbehörde zu bestimmen, der das Forschungsvorhaben anzuzeigen ist und die die Aufgabe hat, die Tätigkeiten und Aufsichtsmaßnahmen aller an sich zuständigen Datenschutzaufsichtsbehörden zu koordinieren. Federführende Datenschutzaufsichtsbehörde sollte die Behörde sein, die für die das Forschungsvorhaben durchführende Stelle zuständig ist.

Dies würde Antragsverfahren beschleunigen, Aufwände der Datenschutzbeauftragten reduzieren und Anreize für die genehmigenden Behörden schaffen, einheitliche und transparente Antrags-/Entscheidungsprozesse zu etablieren. Eine zentralisierte Verwaltung von Datenzugängen und -bereitstellungen würde nicht nur zu einer schnelleren Bereitstellung führen, sondern auch Ressourcen einsparen, die dann für andere Aufgaben beim Datenzugang, wie z.B. die Digitalisierung von Prozessen, eingesetzt werden könnten.[16]

Neben diesen drei Kernforderungen sollte das Forschungsdatengesetz weitere Aspekte regeln, die eine zeitgemäße und forschungsfreundliche Forschungsdateninfrastruktur gestalten.

2. Forschungsgeheimnis, Beschlagnahmeverbot und Zeugnisverweigerungsrecht

Empirische Forschung ist oft auf die unmittelbare Erhebung von Daten bei Personen angewiesen. Aussagekräftige, insbesondere sensible Daten können in der Regel nur unter der Bedingung von Vertraulichkeit erhoben werden. Dies gilt nicht nur die empirische Sozialfor-

14 Eine Ausnahme besteht nur nach § 5 GNDG für dessen Anwendungsbereich.

15 Eine derartige Regelung würde ins Forschungsdatengesetz bzw. für die Statistik ins Bundesstatistikgesetz gehören.

16 Erste Vorschläge in diese Richtung finden sich im GDNG und im Entwurf eines FDG vom 17.12.2024, BT-Drs. 20/14262.

schung. Auch andere Wissenschaftszweige benötigen Vertrauen. Insofern gehört die vertrauliche Datenerhebung zur wissenschaftlichen Methode der Erkenntnisgewinnung, die durch die Forschungsfreiheit nach Art. 5 Abs. 3 Satz 1 GG geschützt ist.[17]

Problematisch ist allerdings, dass das geltende Strafprozessrecht keine expliziten Regelungen zum Schutz der Vertraulichkeit wissenschaftlicher Forschung enthält.[18] Forschende erfahren im Rahmen ihrer Forschungstätigkeit zum Teil intime, aber auch strafrechtlich relevante Informationen, insbesondere wenn dies Gegenstand der Forschungsprojektes ist, wie bspw. im Fall kriminologischer, psychologischer oder soziologischer Feldstudien. Nach der derzeitigen Rechtslage dürfen die Forschenden, da ihnen kein Zeugnisverweigerungsrecht nach §§ 52, 53 StPO zusteht, grundsätzlich[19] als Zeugen vernommen und ihre Studienunterlagen nach § 94 Abs. 1 StPO beschlagnahmt werden.

Der zum Ende der 20. Legislaturperiode ins Parlament eingebrachte Gesetzentwurf eines Forschungsdatengesetzes[20] enthielt in § 9 eine Regelung über das Forschungsgeheimnis, die durch eine daran anschließende Strafvorschrift abgesichert wurde. Für die medizinische Forschung hatte die Datenschutzkonferenz (DSK) bereits mit einer Entschließung im Jahr 2004[21] die Einführung eines Forschungsgeheimnisses gefordert und diese Forderung mit einer weiteren Entschließung im März 2022 bekräftigt.[22] Fast ein Jahrzehnt zuvor war festgestellt worden, dass ein medizinisches Forschungsgeheimnis die „deutschen Forschungsblockaden" beseitigen könnte.[23] Ein zu verabschiedendes Forschungsdatengesetz sollte zwingend eine solche Regelung zu einem allgemeinen, d.h. nicht nur die medizinische Forschung betreffenden Forschungsgeheimnis enthalten.

Die DSK hatte in ihren Entschließungen zur Einführung eines Forschungsgeheimnisses zudem gefordert, die Beschlagnahme von Forschungsdaten zu verbieten und ein entsprechendes Zeugnisverweigerungsrecht zu schaffen.[24] Aufgrund einer polizeilichen Beschlagnahme von Interviewaufzeichnungen bei einem Hochschullehrer wies das BVerfG in seinem Beschluss vom 25.9.2023 darauf hin, dass „erhebliche Bedenken hinsichtlich der Verfassungsmäßigkeit" der Beschlagnahme bestünden.[25] Es betonte, dass gerade die empirische Forschung auf die Verarbeitung von (mitunter sensiblen) Daten angewiesen sei, die oftmals nur unter der Bedingung von Vertraulichkeit erhoben werden könnten. Diese vertrauliche Datenerhebung gehört nach Auffassung des höchsten deutschen Gerichts zur grundrechtlich geschützten wissenschaftlichen Methode.[26]

Dass Forschende aufgrund der Entscheidung des BVerfG vom 25.9.2023 zukünftig auf einen umfassenderen Schutz ihrer Forschungsdaten hoffen können,[27] ist nicht ausreichend. Der Gesetzgeber ist aufgerufen, die durch die unklare Rechtslage entstandenen Unsicherheiten durch klare Regelungen in der Strafprozessordnung zu beseitigen und die §§ 53 und 94 Abs. 1 StPO der Rechtsprechung des BVerfGs anzupassen. Dies im Forschungsdatengesetz zu regeln wäre naheliegend.

3. Reform des Bundesstatistikgesetzes

Eine wichtige und umfangreiche Datenquelle sind die Daten der amtlichen Statistik, die auf Basis des Bundesstatistikgesetzes und der entsprechenden fachstatistischen Einzelgesetze erhoben werden. Um den Forschungsstandort Deutschland zu stärken und der fortschreitenden Digitalisierung Rechnung zu tragen, ist eine kontinuierliche Weiterentwicklung und Anpassung des Bundesstatistikgesetzes erforderlich. Dies betrifft im Wesentlichen drei Themenbereiche: 1) eine stärkere Forschungsorientierung des Statistischen Bundesamtes, 2) eine Modernisierung der Zugangsmöglichkeiten für externe Forschende sowie 3) die Erweiterung der Möglichkeiten der Datenzusammenführung.[28] Zusätzlich sollten zur Stärkung der Krisenresilienz verschiedene Anpassungen vorgenommen werden.

Eigene Forschungsaktivitäten von Datenanbietern wie dem Statistischen Bundesamt tragen wesentlich zur

17 BVerfGE 35, 79 (112); 47, 327 (367); 90, 1 (11 f.); 111, 333 (354); hierzu auch Gärditz, StV 2020, 716 (719).

18 Ausführlich hierzu Gärditz, StV 2020, 716 ff.

19 Eine gesetzliche Ausnahme von diesem Grundsatz besteht, wenn der oder die Forschende gleichzeitig Angehöriger eines der in § 53 Abs. 1 Nr. 3 StPO genannten Berufe ist. Dies gilt insbes. für forschende Mediziner, die sich auf ihre ärztliche Schweigepflicht berufen können.

20 Entwurf eines Gesetzes zum verbesserten Zugang zu und zur Nutzung von Daten für die Forschung, Gesetzentwurf der FDP-Fraktion vom 17.12.2024, BT-Drs. 20/14262.

21 Allerdings eingeschränkt auf die medizinische Forschung, Entschließung der 67. DSK vom 25./26.3.2004, abrufbar unter https://www.bfdi.bund.de/SharedDocs/Downloads/DE/DSK/DSKEntschliessungen/67DSK-EinfuehrungEinesForschungsgeheimnissesFuerMedizinischeDaten.html (zuletzt abgerufen am 8.4.2025)

22 Entschließung der 67. DSK vom 25./26.3.2004, vgl. vorherige Fußnote.

23 *Bochnik*, MedR 1994, 398 ff.

24 Die Diskussion hat allerdings bereits vor 50 Jahren begonnen, vgl. *Eser*, ZRP 1971, 102 ff.

25 BVerfG, ZD 2024, 31 Rn. 12.

26 BVerfG, ZD 2024, 31 Rn. 13.

27 So *Rachut*, jurisPR-ITR 2/2024 Anm. 2.

28 Vgl. Wissenschaftlicher Beratungskreis (WBK) der Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder, Positionspapier: Für zukunftsfähige Forschungsdatenzentren der statistischen Ämter v. 27.2.2025, Mannheim.

Qualitätssicherung der bereitgestellten Daten bei und erhöhen die Qualität der Beratung für Forschende. Eigene Forschung ist daher ein Akkreditierungskriterium des RatSWD für Forschungsdatenzentren.[29] Das BStatG sieht hingegen für das Statistische Bundesamt keinen expliziten Forschungsauftrag vor. Diese Lücke sollte durch eine Änderung im BStatG für das Statistische Bundesamt geschlossen werden, um die Zusammenarbeit zwischen Statistischen Bundesamt und der Wissenschaft zu stärken und die Datenqualität bereits bei der Erhebung zu erhöhen. Die Forschung der amtlichen Statistik wäre komplementär zur Forschung der Datennutzenden.

Modernisierungsbedarf besteht auch bei den technischen Zugängen zu den Mikrodaten. Um den Zugang für unabhängige wissenschaftliche Forschung zeitgemäß zu gestalten und um Aufwände in den Forschungsdatenzentren zu reduzieren, sollte ein Datenfernzugriff auf formal anonymisierte Daten rechtssicher im Bundesstatistikgesetz abgebildet werden. Bereits heute ist es möglich, der Wissenschaft innerhalb speziell abgesicherter Bereiche des Statistischen Bundesamtes und der statistischen Ämter der Länder Zugang zu formal anonymisierten Einzeldaten zu gewähren (§ 16 Abs. 6 BStatG). Strittig ist derzeit die Auslegung des Begriffs „innerhalb speziell abgesicherter Bereiche".[30] Die damit verbundene Rechtsunsicherheit sollte durch eine Anpassung des BStatG dahingehend klargestellt werden, dass auch der virtuelle Bereich ein abgesicherter Bereich ist.

Zudem sollte klargestellt werden, dass sich die Geheimhaltungsverpflichtung, die in § 16 Abs. 6 Nr. 2 BStatG formuliert ist, nur auf Analyseergebnisse bezieht, die den geschützten Raum verlassen. Sie bezieht sich nicht auf die Einsicht formal anonymisierter Einzelangaben, solange diese innerhalb der abgesicherten Bereiche verbleiben und nur durch Verpflichtete nach § 16 Abs. 7 BStatG eingesehen werden. Eine Klarstellung würde bestehende Hemmnisse einzelner Landesämter bei der Bereitstellung formal anonymisierter Daten abbauen und damit sowohl für die Wissenschaft als auch für die Forschungsdatenzentren effizientere Strukturen schaffen.

Um amtliche Mikrodaten nachhaltig bereit stellen zu können, wäre eine gesetzliche Verankerung des Betriebs des FDZ destatis hilfreich. Nach § 3 BStatG gehört es zu den Aufgaben des Statistischen Bundesamtes „Einzelangaben nach Maßgabe dieses Gesetzes oder einer anderen Rechtsvorschrift für wissenschaftliche Zwecke bereitzustellen". Die Aufgabe der Bereitstellung von Einzeldaten für die Wissenschaft ist damit grundsätzlich gesetzlich geregelt, die Form der Bereitstellung jedoch nicht. Das Statistische Bundesamt betreibt hierfür das FDZ destatis, das innerhalb des Statistischen Bundesamts die Interessen der Wissenschaft vertritt. Da ein gesetzlicher Auftrag für den Betrieb eines FDZ nicht vorliegt, ist seitens des Gesetzgebers auch keine ausreichende Ressourcenausstattung vorgesehen. Damit besteht die Gefahr, dass das Statistische Bundesamt in Zeiten schwieriger Haushaltslage die Leistungen des FDZ reduziert. Schon jetzt übersteigt die Nachfrage der Wissenschaft die Ressourcen des FDZ in erheblichem Umfang, so dass Anträge nicht zeitnah bearbeitet werden können und Forschungsprojekte sich zeitlich verzögern. Daher sollte das BStatG im Rahmen eines Forschungsdatengesetzes dahingehend geändert werden, dass das Statistische Bundesamt nicht nur die Aufgabe der Einrichtung und des Betriebs eines FDZs erhält, sondern auch die mit einem FDZ verbundenen Aufgaben festgelegt werden. Entsprechende Regelungen sollten auch für das Forschungsdatenzentrum der Statistischen Ämter der Länder gelten.

Der dritte Themenbereich umfasst eine Erweiterung der bestehenden Regelungen zur Datenzusammenführung, wie bereits in Abschnitt II. 1. c) erläutert wurde. Hier müssen die bestehenden Zusammenführungsregelungen nach § 13a BStatG um personenbezogene Daten erweitert werden. Um (verknüpfte) Daten auch langfristig für die Forschung verfügbar zu machen, sind Löschfristen aufzuheben (Abschnitt II.1.c)).[31]

Eine Anpassung des BStatG im Rahmen des FDG legt nahe, auch Folgerungen aus den Krisen der letzten Jahre aufzugreifen.[32] Politik und Verwaltung haben über Jahrzehnte gelernt, mit den vorhandenen, begrenzten Informationsressourcen zu arbeiten. Dies stößt in Krisensituationen an Grenzen, wenn schnell Informationen benötigt werden, um unter Zeitdruck zielgenaue Entscheidungen treffen zu können. Die Covid-19-Pandemie hat verdeutlicht, wie elementar valide Daten in akuten Krisenlagen sind. Zugleich hat sich gezeigt: Das derzeiti-

29 Vgl. https://www.konsortswd.de/angebote/datenzentren/qualitaetssicherung-zertifizierung/akkreditierung/.

30 *Kühling/Sauerborn* in Kühling, BStatG, § 16 Rn. 61.

31 Für Zusammenführungen von Wirtschaftsstatistiken dürfen Kennnummern nach § 1 Abs. 1 Satz 4 Statistikregistergesetz in den Datensätzen mit den Angaben zu den Erhebungsmerkmalen bis zu 30 Jahre gespeichert werden (§ 13a BStatG). Die Löschfrist sollte daher ersatzlos aus dem BStatG gestrichen werden.

32 Vergleiche den Bericht der Kommission Zukunft Statistik vom 15.1.2024 auf, vgl. https://www.destatis.de/DE/Ueber-uns/Leitung-Organisation/KomZS/abschlussbericht.pdf?__blob=publicationFile&v=5.

ge System der amtlichen Statistik kann nicht flexibel genug auf neuartige Bedarfe und Anforderungen reagieren. Die amtliche Statistik muss deutlich agiler werden.

Um auf aktuelle und potentielle Bedrohungslagen (Pandemie, Naturgefahren, Verteidigungsfall) reagieren zu können, ist empirisch abgesichertes wirtschafts- und sozialwissenschaftliches Wissen über Vulnerabilität, Vorbereitungsstand und Resilienz von Bevölkerung, Behörden und Unternehmen unabdingbar.

Damit die amtliche Statistik agiler werden kann, sollten die rechtlichen Rahmenbedingungen angepasst werden. Die amtliche Statistik in Deutschland wird durch nationale Gesetze und europaweite Verordnungen geregelt, wobei das Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz, BStatG) und die Verordnung (EG) 223/2009 über europäische Statistiken (Europäische Statistikverordnung, VO 223/2009) im Vordergrund stehen. Im Gegensatz zum BStatG, das vornehmlich die Datenerhebung (Inputseite) regelt, konzentriert sich die europäische Gesetzgebung auf die Ergebnisse der Statistik (Outputseite). Wäre der nationale Rechtsrahmen des BStatG stärker outputorientiert, erhielte die Statistik den erforderlichen Spielraum, um neue Methoden zu entwickeln, prozessgenerierte Daten zu nutzen und Modernisierungen umzusetzen, durch die sich Kosten und Belastung reduzieren lassen. Diese Flexibilität benötigt die amtliche Statistik in Deutschland nicht zuletzt, um die von der europäischen Statistik geforderten statistischen Datenbedarfe ordnungsgemäß erfüllen zu können. Eine solche Neuausrichtung der amtlichen Statistik, eine Steigerung der Agilität und eine Öffnung hin zur Wissenschaft und Forschung erfordert eine angemessene Ressourcenausstattung. Vor einer finanziellen Zusatzbelastung der Datennutzenden durch Beiträge kann angesichts der knappen Budgets in wissenschaftlichen Einrichtungen nur gewarnt werden, um prohibitive Wirkungen zu vermeiden.

4. Einrichtung eines Treuhänders für Mikrodaten

Seit über 20 Jahren akkreditiert der RatSWD Forschungsdatenzentren als „Datentreuhänder" für den Zugang der Wissenschaft zu sensiblen Daten. Forschungsdatenzentren ermöglichen der Wissenschaft, eigene Forschungsvorhaben mit vorhandenen Daten durchzuführen. Dies hat den Standort Deutschland für die empirische Sozial- und Wirtschaftsforschung erheblich gestärkt.

Was fehlt, ist die Möglichkeit, die Datenbestände unterschiedlicher Datenproduzenten miteinander zu verknüpfen und damit die vollen Potentiale der Daten zu heben. Genau an diesem Punkt setzt die Forderung nach einem Deutschen Zentrum für Mikrodaten (DZM) als Datentreuhänder an.

Die Aufgabe eines solchen DZM ist es, auf Antrag von Forschenden Datenbestände unterschiedlicher Forschungsdatenzentren miteinander zu verknüpfen und unter Beachtung datenschutzrechtlicher Bestimmungen zur Verfügung zu stellen. In diesem Sinne agiert das DZM als technische Verbindungsstelle der verschiedenen Forschungsdatenzentren und soll Forschungsdatenzentren ergänzen, aber nicht ersetzen. Die Daten der Forschungsdatenzentren erfordern Spezialwissen, welches ein DZM in der Gesamtheit nicht bereitstellen kann; dies betrifft Wissen aus der Datenproduktion und -dokumentation, sowie daraus abgeleitetes und für die Beratung von Forschenden unabdingbares Wissen. Es ist nicht sinnvoll, dieses Spezialwissen bei einem Treuhänder zu bündeln und damit zu doppeln. Das DZM sollte daher die technische Datenverknüpfung und die datenschutzkonforme Bereitstellung dieser Art von Mikrodaten umsetzen und dabei eng mit den FDZ zusammenarbeiten und deren Expertise zu den spezifischen Datenbeständen bei der Verknüpfung nutzen. Ein interdisziplinärer Beirat aus Forschenden und Datenproduzenten sollte die Arbeit des DZM unterstützen und begleiten und seine Expertise in den Aufbau des DZM einbringen. Forschungsprojekte, die ausschließlich Mikrodaten eines spezifischen FDZ nutzen, werden weiterhin über das FDZ ohne Beteiligung des DZM abgewickelt.

Die Einrichtung eines DZM über das Forschungsdatengesetz schließt eine wichtige Lücke und stärkt so den Wissenschaftsstandort Deutschland. Wichtig ist, dass keine Doppelstrukturen aufgebaut werden und es eine Ergänzung des Erfolgsmodells der Forschungsdatenzentren darstellt.

III. Deutschland in europäischen Forschungsdatenräumen

Nach Art. 3 Abs. 3 UAbs. 3 Satz 3 EU-Vertrag ist es das Ziel der Europäischen Union, den wissenschaftlichen und technischen Fortschritt zu fördern. Dies will sie erreichen, indem sie nach Art. 179 Abs. 1 AEUV einen europäischen Forschungsraum geschaffen hat, in dem Freizügigkeit für Forschende herrscht und wissenschaftliche Erkenntnisse und Technologien frei ausgetauscht werden. Dies geht auf eine Entschließung des Europäischen Rates vom 15.6.2000 zurück, mit der ein Europäi-

scher Forschungsraum geschaffen wurde.[33] Der Europäische Rat hat die Bedeutung des Europäischen Forschungsraumes nochmals in seinen „Schlussfolgerungen zum Thema „Neuer Europäischer Forschungsraum" vom 1.12.2020 betont.[34] Darin definiert er den „Neuen Europäischen Forschungsraum" als einen „forschungsorientierten, wertebasierten Bereich, der auf Exzellenz und Wirkung ausgerichtet ist und in dem Forschende und Technologie unterstützt werden und frei zirkulieren können." Es wurde nicht nur die besondere Bedeutung der Grundlagenforschung betont, sondern die Mitgliedstaaten aufgefordert, Investitionen in den „Neuen Europäischen Forschungsraum" Vorrang einzuräumen. Zudem wurde in der von allen Mitgliedstaaten unterzeichneten „Bonner Erklärung zur Forschungsfreiheit" eine gemeinsame Definition der Forschungsfreiheit gefunden und gefordert, europaweite Forschungsdateninfrastrukturen zu schaffen.[35] Dem ist Deutschland mit dem Aufbau der interdisziplinären Nationalen Forschungsdateninfrastruktur (NFDI) mit ihren mittlerweile 27 Fachkonsortien und fünf Sektionen nachgekommen.[36] Ein Forschungsdatengesetz kann diesen Prozess vorantreiben, damit Deutschland seinen Beitrag zur Schaffung europäischer Forschungsdatenräume leisten kann.

Die EU-Kommission hat sich mit der „Europäischen Datenstrategie"[37] das Ziel gesetzt, einen europäischen Binnenmarkt für Daten zu schaffen.[38]

Im Rahmen der Europäischen Datenstrategie hat die EU den Data Governance Act (DGA), den Data Market Act (DMA), den Data Service Act und den Data Act erlassen, die ebenfalls das Ziel verfolgen, für unterschiedliche Zwecke den Zugang zu Daten innerhalb der EU zu erleichtern. Insbesondere der DGA sieht Mechanismen für die Nutzung kommerzieller und nichtkommerzieller Daten innerhalb der EU für unternehmerische als auch für wissenschaftliche Zwecke vor.

Im Rahmen der Datenstrategie hat die EU-Kommission am 23.2.2022 ein „Arbeitsdokument der Kommissionsdienststellen über Gemeinsame Europäische Datenräume" („Common Data Spaces") vorgelegt.[39] Darin hat die EU-Kommission zunächst zwölf Bereiche benannt, in denen sog. Datenräume entstehen sollen.[40] Mit dem Arbeitsdokument vom 24.1.2024[41] wurden weitere Bereiche genannt, so dass die Kommission derzeit 14 sog. Datenräume plant. Mit der EHDS-VO vom 11.2.2025[42] hat die EU mit dem Europäischen Gesundheitsdatenraum den ersten dieser gemeinsamen Datenräume rechtlich geregelt. Mit der Verordnung reguliert die EU den Umgang mit Gesundheitsdaten einerseits für Behandlungszwecke.[43] Die EHDS-VO spricht insoweit von Primärnutzung.[44] Darüber hinaus regelt die EHDS-VO in „Kapitel IV – Sekundärnutzung"[45] u.a. die Nutzung von Gesundheitsdaten für Forschungszwecke (Art. 53 Abs. 1 lit. e EDHS-VO), allerdings eingeschränkt auf die *„wissenschaftliche Forschung im Bereich des Gesundheitswesens oder des Pflegesektors, die zur öffentlichen Gesundheit oder zur Bewertung von Gesundheitstechnologien beiträgt oder ein hohes Maß an Qualität und Sicherheit der Gesundheitsversorgung, von Arzneimitteln oder Medizinprodukten sicherstellt, mit dem Ziel, Endnutzern wie Patienten, Angehörigen der Gesundheitsberufe und Gesundheitsverwaltungen zugutezukommen".* Mit der EHDS-VO verspricht man sich einen „Paradigmenwechsel in der

33 Entschließung des Rates vom 15.5.2000 zur Schaffung eines Europäischen Raums der Forschung und Innovation, ABl. C 205 v. 19.7.2000, S. 1.

34 „Schlussfolgerungen des Rates zum Thema neuer Europäischer Forschungsraum" vom 1.12.2020, 13567/20, abrufbar unter https://data.consilium.europa.eu/doc/document/ST-13567-2020-INIT/de/pdf (zuletzt abgerufen am 3.02.2025).

35 „Bonner Erklärung zur Forschungsfreiheit" anlässlich der Ministerkonferenz zum Europäischen Forschungsraum am 20.10.2020 in Bonn, abrufbar unter https://www.bmbf.de/SharedDocs/Downloads/DE/b/bonner_erklaerung_deu.pdf?__blob=publicationFile&v=4 (zuletzt abgerufen am 3.2.2025).

36 Ausführlich zum Europäischen Forschungsraum und zur NFDI, vgl. Raum in *Ehmann/Selmayr*, DSGVO, Art. 89 Rn. 6.

37 Mitteilung der Kommission an das Europäische Parlament, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen – Eine europäische Datenstrategie – vom 19.2.2020, COM(2020) 66 final.

38 *Klumpen/Schliffka/Polus*, WISTA 2021 (Heft 6), 22.

39 Commission Staff Working Document on Common European Data Spaces on 23.2.2022, SWD(2022) 45 final.

40 Ausführlich zu den Europäischen Datenräumen, vgl. *Raum* in Ehmann/Selmayr, DSGVO, Art. 89 Rn. 6.

41 Commission Staff Working Document on Common European Data Spaces on 24.1.2024, SWD(2024) 21 final.

42 Verordnung (EU) 2025/327 des Europäischen Parlaments und des Rates über den europäischen Gesundheitsdatenraum sowie zur Änderung der Richtlinie 2011/24/EU und der Verordnung (EU) 2024/2847 vom 11.2.2025, ABl.EU vom 5.3.2025.

43 COM(2022) 197 final, S. 1 ff; vgl. auch *Buchholtz/Schmalhorst/Brauneck*, MDR 2024, 171 (172).

44 Art. 2 Abs. 2 lit. d) EHDS-VO definiert „Primärnutzung" als „die Verarbeitung elektronischer Gesundheitsdaten für die Gesundheitsversorgung zur Beurteilung, Erhaltung oder Wiederherstellung des Gesundheitszustands der natürlichen Person, auf die sich diese Daten beziehen, einschließlich der Verschreibung, Abgabe und Bereitstellung von Arzneimitteln und Medizinprodukten, sowie für die einschlägigen Sozialleistungen und für die einschlägigen Dienste der Verwaltung oder Kostenerstattung".

45 Art. 2 Abs. 2 lit. e) EHDS-VO versteht die Verordnung „Sekundärnutzung" als „die Verarbeitung elektronischer Gesundheitsdaten für die in Kapitel IV der vorliegenden Verordnung genannten Zwecke, sofern es sich nicht um die Zwecke handelt, für die sie ursprünglich erhoben oder erstellt wurden".

Gesundheitsdatennutzung".[46] Im internationalen Bereich zeigt sich, dass eine Nutzung von Gesundheitsdaten für die medizinische Forschung zu einer Verbesserung der Gesundheitsversorgung führen kann.[47]

Im Vorgriff auf die EHDS-VO[48] hat der deutsche Gesetzgeber mit dem GDNG vom 25.3.2024[49] eine gesetzliche Regelung erlassen, die eng auf einen überschaubaren Bereich der medizinischen Forschung beschränkt ist. Ziel des GDNG ist es, bei gleichzeitiger Stärkung des Gesundheitsdatenschutzes die Verknüpfbarkeit von Gesundheitsdaten zu erleichtern.[50] Dies soll u.a. dadurch gelingen, dass das GDNG die Möglichkeit einräumt, die in der sog. elektronischen Patientenakte[51] verfügbaren Daten für Forschungszwecke zu nutzen. Nach § 4 des derzeitigen GDNG gilt dies jedoch nur für die Verknüpfung von Daten aus dem FDZ Gesundheit mit Daten der klinischen Krebsregister der Länder.[52] Die Verknüpfungen von Daten aus dem FDZ Gesundheit mit den teilweise älteren und damit umfangreicheren epidemiologischen Krebsregistern der Länder[53] und dem Zentrum für Krebsregisterdaten (ZfKD) im Robert Koch-Institut sind damit bereits nicht zulässig. Die Möglichkeit der Verknüpfung mit anderen Daten, um berufsbedingte Erkrankungen zu erforschen oder soziologische Daten mit Gesundheitsdaten zu verknüpfen, ist noch nicht vorgesehen, sollte aber durch das Forschungsdatengesetz geregelt werden.

Diese Rechtsetzung des europäischen und nationalen Gesetzgebers mag als wichtiger Schritt für die Nutzung von Gesundheitsdaten zu Forschungszwecken gefeiert werden. Es besteht allerdings die Gefahr, Datensilos für jeweils enge Forschungszwecke zu schaffen. Bedauerlicherweise ist mit der EHDS-VO ein Datensilo für medizinische Daten im Entstehen. Sollte die EHDS-VO als Blaupause für die anderen EU-weiten Datenräume dienen, ist zu befürchten, dass auch hier weitere isolierte Datensilos entstehen werden. Eine Verknüpfungsmöglichkeit der Daten zwischen diesen isolierten Datensilos wird nicht mitgedacht. Im Vordergrund steht jeweils der Vorteil für die enge Forschung mit den für die Daten unmittelbar naheliegenden Forschungszwecken.[54] Gesundheits-, Mobilitäts-, Bildungs- oder Arbeitsmarktdaten sind aber auch in jeweils weiteren Forschungsbereichen für die wissenschaftliche Forschung unverzichtbar. Es reicht daher nicht aus, bereichsbezogen die Nutzung, insbesondere die Verknüpfbarkeit, von Daten für bereichspezifische Forschungszwecke zu regeln. Vielmehr bedarf es allgemeiner Regelungen dazu, unter welchen Voraussetzungen Daten zu offenen Forschungszwecken verwendet und insbesondere miteinander verknüpft werden können. Dem FDG kommt die Aufgabe zu, diese allgemeinen Regeln zu schaffen, damit keine neuen „Datensilos" entstehen, deren Nutzen für die Wissenschaft begrenzt ist.

IV. Weitere drängende Themen

Es gibt Desiderate der deutschen Forschungsdateninfrastruktur, die über die unmittelbaren Kernforderungen an ein Forschungsdatengesetz hinausgehen. Wir diskutieren die Bedeutung eines Bildungsverlaufsregisters sowie die Datenkooperation zwischen Unternehmen und Wissenschaft.

1. Datenlücken schließen: Umsetzung eines Bildungsverlaufsregisters

Das FDG muss die wissenschaftliche Nutzung bereits vorhandener Daten regeln. Es ist so zu formulieren, dass noch zu schließende Datenlücken berücksichtigt werden. Erhebliche Datenlücken bestehen im Bildungsbereich, was angesichts der zentralen Bedeutung von Bildung für die gesellschaftliche und wirtschaftliche Ent-

46 Vgl. Entschließung des Bundesrates zur Ausgestaltung des Gesundheitsdatennutzungsgesetzes vom 16.12.2022, BR-Dr. 597/22 (Beschluss), S. 3.

47 Dies gilt für die Gesundheitssysteme in Finnland, Island oder Israel, vgl. *Rauch/Richters/Naucke*, GesR 2024, 218 (219) und Diercks, PharmR 2023, 369 (370). Dierks weist insbes. auf das finnische Gesetz zur Sekundärnutzung von Sozial- und Gesundheitsdaten (sog.Toisiolaki – Laki sosiaali- ja terveystietojen toissijaisesta käytöstä), das als Modell für die EHDS-VO gestanden hat, vgl. Diercks, PharmR 2023, 369 (373).

48 BT-Dr. 20/9046, S. 3 f.

49 Vgl. Fn. 4.

50 BT-Dr. 20/9046, S. 3. Die Aussage hört sich wie eine Quadratur des Kreises an.

51 Dabei handelt es sich tatsächlich um eine von den gesetzlichen Krankenkassen bereitgestellte elektronische Gesundheitsakte. Bürgerinnen und Bürger sind bei den Krankenkassen Versicherte, keine Patientinnen und Patienten.

52 Klinische Krebsregister werden in den Ländern verbindlich erst durch § 65 SGB V.

53 Epidemiologische Krebsregister gab es auf deutschem Boden zunächst nur in der ehemaligen DDR, im Saarland und an der Universität Mainz (Kinderkrebsregister). Ab den 1990er Jahren wurden dann in allen Bundesländern epidemiologische Krebsregister eingerichtet. Das Krebsregister der ehemaligen DDR wurde als gemeinsames Krebsregister der neuen Bundesländer und des Landes Berlin weitergeführt.

54 Die EU-Kommission spricht davon, dass dieses „gesundheitsspezifische (…) Ökosystem aus Vorschriften, gemeinsamen Standards und Verfahren, Infrastrukturen und einem Governance Rahmen" bestehen soll, COM (2022) 196, S. 10.

wicklung und damit für den Wohlstand Deutschlands kaum hinnehmbar ist.[55] Spätestens seit dem PISA-Schock im Jahr 2001 ist in Deutschland bekannt und unstrittig, dass das deutsche Bildungssystem weiterentwickelt werden muss. Dazu bedarf es empirischer Evidenz. Der Auftrag zum Bildungsmonitoring ist mit Art. 91b Satz 2 GG bereits im Grundgesetz verankert.[56] Ohne eine gute Datenbasis kann aber weder ein gutes Bildungsmonitoring noch ein evidenzbasiertes Handeln von Politik und Verwaltung gelingen.[57]

Obwohl die große Bedeutung von Daten für eine gute Bildungspolitik seit vielen Jahren bekannt ist und in Deutschland seit geraumer Zeit über ein Bildungsverlaufsregister diskutiert wird,[58] ist die Datenbasis in Deutschland unzureichend. Der 2003 von der Kultusministerkonferenz (KMK) beschlossene Schülerkerndatensatz ist noch nicht umgesetzt.[59] Daten zur Soziodemografie der Schülerschaft oder Erfolgsindikatoren wie Übergangs- und Abschlussquoten werden in der amtlichen Statistik der meisten Bundesländer nach wie vor nur in aggregierter Form bereitgestellt. Und dort, wo individuelle amtliche Bildungsverlaufsdaten vorliegen, ist der Zugang der Wissenschaft zu den vorhandenen amtlichen Mikrodaten häufig mit hohen Hürden verbunden. In Deutschland sind Bildungsverläufe häufig nur auf der Basis von Befragungsdaten wie dem Nationalen Bildungspanel (NEPS) möglich.[60] Befragungsdaten alleine reichen jedoch für eine effektive und evidenzbasierte Bildungspolitik nicht aus. Erforderlich ist ein breit angelegtes Bildungsverlaufsregister, das alle Bildungsabschnitte von der frühkindlichen Bildung über den Primar- und Sekundarbereich bis hin zur beruflichen und akademischen Bildung umfasst. Auch das lebenslange Lernen durch betriebliche und außerbetriebliche Weiterbildung sollte erfasst werden. Zudem lassen sich Bildungserfolge nur unzureichend über Bildungsübergänge und erreichte Abschlüsse abbilden. Für eine effektive Steuerung des Bildungswesens sind standardisierte, regelmäßige und flächendeckende Kompetenzmessungen unerlässlich, die mit den Individualdaten eines Kerndatensatzes verknüpft werden müssten. Der Koalitionsvertrag von Union und SPD[61] für die 21. Legislaturperiode stellt den Aufbau eines Bildungsverlaufsregisters in Zusammenarbeit mit den Ländern in Aussicht. Nun kommt es auf die konkrete Ausgestaltung an, an der die Wissenschaft schon frühzeitig bei Konzeption und Aufbau einbezogen werden sollte.

Die beschriebenen Defizite erschweren eine zielgerichtete, effektive und effiziente Bildungspolitik und stellen einen erheblichen Wettbewerbsnachteil für den Forschungsstandort Deutschland dar. So können beispielsweise keine Aussagen über die Bildungs- und Erwerbsbiografien von Schulabgängern getroffen werden, weil die notwendigen Daten fehlen. Mit einem für die Wissenschaft nutzbaren Bildungsverlaufsregister stünde eine Datenbasis zur Beantwortung dieser Fragen zur Verfügung.

Bildungsverlaufsregister sind in vielen europäischen Ländern längst Realität. So stellen u.a. die skandinavischen Länder, die Niederlande, Österreich, die Schweiz und Großbritannien umfassende und flächendeckende Bildungsverlaufsdaten für die Bildungsforschung und -steuerung zur Verfügung. Aus den Infrastrukturen der europäischen Nachbarländer lassen sich Anregungen für die institutionelle und datenschutzkonforme Ausgestaltung einer entsprechenden Infrastruktur in Deutschland ableiten.

Bildungsdaten dürfen aber nicht zu einem neuen Datensilo werden. Für viele Fragestellungen von hoher gesellschaftlicher Relevanz ist eine Verknüpfbarkeit der Bildungsdaten mit anderen Daten erforderlich, denn Bildungserfolg bemisst sich nicht nur an Kompetenzmessungen und Bildungsabschlüssen. Gute Bildungspolitik kann zugleich Arbeitsmarkt-, Wirtschafts-, Sozial- und Familienpolitik sein, und bessere Bildung und weniger Bildungsungleichheit können erhebliche Auswirkungen auf Armut, Gesundheit, Kriminalität und vieles mehr haben. Daher muss das Bildungsverlaufsregister mit Datenregistern aus anderen gesellschaftlichen Bereichen, wie Arbeitsmarkt- und Gesundheitsdaten, ver-

55 Vgl. *Hertweck* et al. (2023). Bildungsdaten: Datenlücken durch ein Bildungsverlaufsregister schließen, Wirtschaftsdienst 103(11), 733-736.

56 Artikel 91b Satz 2 GG: „(2) Bund und Länder können auf Grund von Vereinbarungen zur Feststellung der Leistungsfähigkeit des Bildungswesens im internationalen Vergleich und bei diesbezüglichen Berichten und Empfehlungen zusammenwirken."

57 Vgl. auch SWK, Entwicklung von Leitlinien für das Monitoring und die Evaluation von Förderprogrammen im Bildungsbereich. Berlin 2022, abrufbar unter https://www.kmk.org/fileadmin/Dateien/pdf/KMK/SWK/2022/SWK-2022-Impulspapier_Monitoring.pdf (zuletzt abgerufen am 08.04.2025).

58 Vgl. RatSWD (2022), RatSWD fordert ein Bildungsverlaufsregister, abrufbar unter https://www.konsortswd.de/wp-content/uploads/Positionspapier-RatSWD-Aufbau-eines-Bildungsverlaufsregisters.pdf

59 Vgl. KMK, Ländervereinbarung über die gemeinsame Grundstruktur des Schulwesens und die gesamtstaatliche Verantwortung der Länder in zentralen bildungspolitischen Fragen v. 15.10.2020, abrufbar unter https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2020/2020_10_15-Laendervereinbarung.pdf

60 https://www.neps-data.de/.

61 https://www.cdu.de/app/uploads/2025/04/Koalitionsvertrag-2025-1.pdf.

knüpft werden können. Ebenso sinnvoll ist die Möglichkeit der Verknüpfung mit anderen umfragebasierten Datensätzen, um Lücken in den Registerdaten gezielt für bestimmte Forschungsfragen schließen zu können.

2. Prozessdaten von Unternehmen nutzbar machen

Der vorliegende Entwurf des Forschungsdatengesetzes bezieht sich auf Verwaltungs- und Registerdaten sowie Daten, die in Forschungsdatenzentren vorliegen. Daten von Unternehmen, die in Geschäftsprozessen entstehen, werden nicht adressiert. Daten von Unternehmen können jedoch eine wertvolle Grundlage für Forschung sein. Beispiele hierfür sind Nutzungsdaten von sozialen Medien, Transaktionsdaten von Finanzdienstleistern oder Einzelhandelsgeschäften, Mobilitätsdaten, Telekommunikationsdaten oder Personaldaten aus HR-Prozessen.

Das Teilen von Daten von Unternehmen mit der Wissenschaft (Business to Science – B2S) erfordert geeignete und innovative Kooperationsmodelle zwischen Unternehmen und der Wissenschaft. Dies ist herausfordernd, wenn die Vorteile einer Kooperation auf den ersten Blick eher auf Seiten der Wissenschaft verortet werden als auf Seiten der kooperierenden Unternehmen.[62] Einen gesetzlichen Zwang, dass alle Unternehmen gleichermaßen alle ihre Daten der Wissenschaft kostenfrei zur Verfügung stellen müssen, sollte es nicht geben.[63] Jedoch könnte eine im FDG angelegte Förderung für die Entwicklung von B2S-Kooperationsmodellen entsprechende Entwicklungen in diese Richtung unterstützen. Damit die deutsche Forschung mit der internationalen Forschung (insbesondere für die USA) mithalten kann, sind Kooperationsmöglichkeiten erforderlich. Um bessere Zugangswege für die Wissenschaft zu Daten von Unternehmen auszuloten, gibt es bereits die NFDI Sektion "Industry Engagement", die Stifterverbandsinitiative "Datagroup Business 2 Science" und das Projekt "Zugang zu Firmendaten für die Wirtschaftswissenschaften" des NFDI-Konsortiums "KonsortSWD".

In der EU ist seit 2024 das Datengesetz (Verordnung (EU) 2023/2854) der rechtliche Rahmen, in dem die Potentiale von Daten besser genutzt werden und gleichzeitig Daten geschützt werden sollen. Das Datengesetz regelt den Umgang mit Daten im digitalen Zeitalter, insbesondere mit Daten, die bei der Nutzung vernetzter Produkte (z. B. vernetzte Autos, Medizin- und Fitnessgeräte, Industrie- oder Landmaschinen) anfallen. Das faire Datenteilen zwischen Unternehmen soll gefördert werden (B2B), es ist sogar obligatorisches Datenteilen gegen eine angemessene Entschädigung vorgesehen (Kapitel II). Die Europäische Kommission wird dazu Musterverträge erarbeiten und empfehlen.[64] Öffentliche Stellen bekommen zudem Zugang zu Daten von Unternehmen (B2G), insbesondere, wenn diese in Krisensituationen erforderlich sind (Kapitel V). Ziel des Datengesetzes ist die Stärkung der Datenwirtschaft der EU und die Förderung eines wettbewerbsfähigen Datenmarkts. Es gibt im Datengesetz jedoch keine Bestimmung, die Unternehmen explizit dazu verpflichtet, ihre Daten mit der Wissenschaft zu teilen. Vielmehr geht es darum, gleiche Wettbewerbsbedingungen zu schaffen und sicherzustellen, dass schutzwürdige Geschäftsinteressen (wie Betriebs- und Geschäftsgeheimnisse) gewahrt bleiben. Darüber hinaus gibt es ergänzende Regelungen, z.B. im Rahmen des Daten-Governance-Gesetzes, die unter dem Begriff „Datenaltruismus" den freiwilligen Datenaustausch zum Wohle der Allgemeinheit – wozu auch wissenschaftliche Forschung zählen kann – fördern sollen.[65]

Dass Datenkooperationen wünschenswert sind, ist offensichtlich. Unternehmensdaten können Lücken im bestehenden Datenangebot der amtlichen Statistik und der zur Verfügung stehenden Surveydaten schließen. Es gibt bereits Beispiele dafür, wie Unternehmensdaten für die Wissenschaft in FDZ zur Verfügung gestellt werden können. So stehen über das FDZ Ruhr kleinräumige Daten[66] und Daten aus der Kooperation mit einer Immobilienplattform[67] für die wissenschaftliche Nutzung zur Verfügung. Diese Kooperationen sind Beispiele dafür, dass und wie Daten aus Unternehmenskooperationen für die Forschung über ein FDZ sicher verfügbar gemacht werden können.

62 Vgl. *Stahl, Hamann, Hoff, Kockelmann*, Collaboration Models Between Industry and Academia. Whitepaper on behalf of the Section Industry Engagement of the National Research Data Infrastructure (NFDI), 2024, abrufbar unter https://zenodo.org/records/10473579 (zuletzt abgerufen am 8.4.2025).

63 Einige marktbeherrschende (Digital-)Unternehmen sind schon jetzt gesetzlich zur Datenbereitstellung verpflichtet und weitere derartige Verpflichtungen könnten sinnvoll sein, um negative Auswirkungen auf Wettbewerb und die Gesellschaft unabhängig zu untersuchen.

64 Verordnung (EU) 2023/2854 des Europäischen Parlaments und des Rates vom 13.12.2023 über harmonisierte Vorschriften für einen fairen Datenzugang und eine faire Datennutzung sowie zur Änderung der Verordnung (EU) 2017/2394 und der Richtlinie (EU) 2020/1828 (Datenverordnung) [Data Act] vom 13.12.2023, ABl.EU vom 22.12.2023.

65 Es wäre wünschenswert, dass die Bundesregierung ihrer seit dem 24.9.2023 überfälligen vertraglichen Verpflichtung, das Daten-Governance-Gesetz in Deutschland umzusetzen, nachkommt.

66 Vgl. RWI, Sozioökonomische Daten auf Rasterebene - Scientific Use File (Welle 14), abrufbar unter DOI: 10.7807/microm:suf:v14.

67 Vgl. RWI, RWI Real Estate Data - Wohnungskauf – SUF, abrufbar unter DOI: 10.7807/immo:red:wk:suf:v11

Der Zugang der Forschung zu diesen aus Forschungssicht sehr wertvollen Daten beruht in der Regel meist auf projektspezifischen Unternehmenskooperationen, die in Deutschland aufgrund des hohen administrativen Aufwands und der rechtlichen Unsicherheit noch selten sind. Zudem werden diese Daten (in anonymisierter oder aufbereiteter Form) häufig von den Unternehmen selbst oder kommerziellen Datenhändlern zum Kauf angeboten – oft zu für die unabhängige Forschung prohibitiv hohen Preisen und mit restriktiven Lizenzbedingungen. Das FDG könnte hier ansetzen und zum einen Rechtssicherheit für datenschutzkonformes Teilen von Daten im B2S-Bereich (analog zum B2B-Bereich im Data Act) schaffen, um auch für Unternehmen das Datenteilen attraktiv zu machen. Zum anderen könnte überlegt werden, wie der Zugang zu kommerziell vermarkteten Daten für nichtkommerzielle Forschung ermöglicht werden kann, beispielsweise über das Einbinden eines Datentreuhänders.

V. Fazit

Der Zugang zu Forschungsdaten in Deutschland ist im internationalen Vergleich rückständig. Darunter leiden nicht nur Forschende. Mangelhafte Daten als Entscheidungsgrundlage führen zu ineffektiver und ineffizienter Politik. Das verursacht hohe gesellschaftliche Kosten. Viele relevante Forschungsfragen können aufgrund der schlechten Forschungsdatenbasis nicht beantwortet werden. Das schadet Deutschland als Forschungsstandort. Die europäischen Nachbarländer Deutschlands zeigen mit ihrer modernen Forschungsdatenpolitik die Möglichkeiten auf, die in aktuellen und detaillierten Informationen stecken.

Ein modernes Forschungsdatengesetz sollte daher die Potentiale des bestehenden deutschen Datenökosystems für Wissenschaft, Wirtschaft und Gesellschaft heben, indem es Datensilos aufbricht und Datenverknüpfung ermöglicht. Kurzfristig müssen Fragen des Datenzugangs, der Datenverknüpfbarkeit, des Datenschutzes und der gesetzlichen Vorgaben für die amtliche Statistik angegangen werden. Für die Umsetzung der Datenverknüpfung ist ein Datentreuhänder mit klar definierten Zuständigkeiten und hinreichender Ressourcenausstattung zu etablieren.

Ein Forschungsdatengesetz muss aber über die Nutzung bestehender Daten hinaus die Weiterentwicklung des Datenökosystems unterstützen. So ist der Aufbau eines nationalen Bildungsverlaufsregisters zur Stärkung eines chancengerechten und leistungsfähigen Bildungssystems überfällig. Das Bildungsverlaufsregister unterstützt die Umsetzung transparenter, zielgerichteter und evidenzbasierter Bildungspolitik. Daten von Unternehmen sind eine weitere Datenquelle, der perspektivisch eine größer werdende Bedeutung zukommt. Das bislang in Deutschland vernachlässigte Potenzial der Datenkooperation zwischen Unternehmen und Forschung sollte durch die Entwicklung von Kooperationsmodellen belebt werden. Auch diese Aspekte können im FDG verankert werden.

Parallel zur Weiterentwicklung der nationalen Regelungen für Forschungsdaten ist es erforderlich, einen Beitrag zur Diskussion um europäische Datenräume zu leisten. Die dort angestoßene Entwicklung geht da in die falsche Richtung, wo sie sich auf die Schaffung von Datensilos mit engen Zweckbindungen fokussiert. Während die Initiativen zur Bereitstellung von Forschungsdaten grundsätzlich zu begrüßen sind, sollte die Bundesregierung ihren Einfluss nutzen, um die Fehler, die sich bereits im engen Anwendungsbereich des GDNG in Deutschland spiegeln, in anderen Themenbereichen zu vermeiden. Breite Verknüpfungsmöglichkeiten müssen von Anfang an mitgedacht und die Bildung von Silos verhindert werden.

Die Rechtsmaterie, die von einem FDG angesprochen wird, umfasst viele Regelungsbereiche und ist durch die Überlagerung der zuständigen Gebietskörperschaften (EU, Bund, Länder) und der Heterogenität der betroffenen Rechtskreise komplex. Die Natur des Problems ruft nach einer großen Lösung.

Prof. Regina T. Riphahn, Ph.D. ist Professorin für Statistik und empirische Wirtschaftsforschung an der Friedrich-Alexander-Universität Erlangen-Nürnberg. Sie war von 2008 bis 2014 Mitglied des Wissenschaftsrates und ab 2012 Vorsitzende seiner wissenschaftlichen Kommission. Von 2014 bis 2020 war sie Vorsitzende des Rates für Sozial- und Wirtschaftsdaten (RatSWD) und von 2023 bis 2024 Vorsitzende des Vereins für Socialpolitik. Seit 2017 ist sie Vizepräsidentin der Nationalen Akademie der Wissenschaften Leopoldina.

Prof. Dr. Andreas Peichl ist Leiter des ifo Zentrums für Makroökonomik und Befragungen und Professor für Volkswirtschaftslehre, insb. Makroökonomie und Finanzwissenschaft an der Ludwig-Maximilians-Universität München. Er war und ist an zahlreichen Forschungsprojekten für nationale Ministerien, die Europäische Kommission, das Europäische Parlament und die OECD beteiligt. Er ist unter anderem Mitglied des Wissenschaftlichen Beirats des Bundesfinanzmi-

nisteriums und war Mitglied in der Gründungskommission des Dateninstituts der Bundesregierung.

Rechtsanwalt MinR a.D. Dr. Bertram Raum leitete zwischen 2009 und 2022 beim BfDI das Referat „Datenschutz im Sozial- und Gesundheitswesen". Seit 2023 ist er Rechtsanwalt und seit Anfang 2025 Gastforscher mit dem Schwerpunkt „Digitalisierung im Gesundheitswesen" beim Institut für Digitalisierung der Rechtswissenschaftlichen Fakultät der Universität zu Köln.

Prof. Dr. Kerstin Schneider ist Professorin für Finanzwissenschaft und Steuerlehre der Fakultät für Wirtschaftswissenschaft – Schumpeter School of Business and Economics der Universität Wuppertal und Mitglied im erweiterten Vorstand des RWI – Leibniz Institut für Wirtschaftsforschung. Von 2020 bis 2024 war sie Stellvertretende Vorsitzende des Rates für Sozial- und Wirtschaftsdaten (RatSWD), seit 2024 ist sie die Vorsitzende.

Dr. Daniel Vorgrimler ist seit Juli 2019 Leiter der Abteilung „Strategie und Planung, Internationale, Forschung und Kommunikation" des Statistischen Bundesamtes. Seit 2020 Mitglied im Rat für Sozial- und Wirtschaftsdaten (RatSWD) und seit 2024 dessen Stellvertretender Vorsitzender. Er hat Lehraufträge an der Universität Mainz und Hochschule Fresenius.

Das Autorenteam dankt Thomas Bauer, Hendrik Jürges, Tatjana Mika, Sebastian Wichert und Joachim Winter für hilfreiche Kommentare.