

Juni 2023

Diskussionsimpuls: Anforderungen an die Ausgestaltung eines Forschungsdatengesetzes und eines Dateninstituts

Der RfII hat sich wiederholt für einen durch Gesetz privilegierten Zugang der öffentlich getragenen Wissenschaft zu Daten der öffentlichen Hand sowie aus Wirtschaft und Gesellschaft ausgesprochen. Ein solches Forschungsprivileg ist nichts Neues. Es hat in der analogen Welt eine lange Tradition und korrespondiert direkt mit den spezifischen Leistungen, die das öffentliche Wissenschaftssystem durch seine Beiträge im Innovationssystem, in der Daseins- und Zukunftsvorsorge sowie bei der Lösung großer globaler Herausforderungen für die Gesellschaft erbringt. Allerdings fehlt dem Forschungsprivileg noch eine hinreichende Ausarbeitung, die den sich ändernden gesellschaftlichen Bedingungen durch die digitale Transformation Rechnung trägt. Insofern begrüßt der RfII politische Anstrengungen zur Verbesserung des Zugangs der öffentlich getragenen Wissenschaft zu Daten, die in anderen gesellschaftlichen Sektoren, also auch außerhalb der Wissenschaft, erhoben bzw. generiert werden. Der RfII hält darüber hinaus auch Regulierungen des externen Zugriffs auf Forschungsdaten für sinnvoll, die aktuell im Fokus umfassender Bemühungen zur Gestaltung des digitalen europäischen Binnenmarkts und der digitalen Transformation nationaler Innovationssysteme stehen. Die deutsche Bundesregierung verfolgt in diesem Zusammenhang und eingerahmt von den bereits erfolgten und noch laufenden datenbezogenen Gesetzgebungsprozessen auf der europäischen Ebene zwei wichtige Gestaltungsvorhaben: die Einführung eines Forschungsdatengesetzes und den Aufbau eines Dateninstituts in Deutschland.

AUSGANGSFRAGEN

Die Ausgestaltung eines **Forschungsdatengesetzes** und die Einrichtung eines **Dateninstituts** sind zentrale Vorhaben der Bundesregierung, die das Teilen von Daten und den Zugang zu Daten im Interesse der Wissenschaft entscheidend verbessern können. Die Vorhaben werden unter unterschiedlichen Ressortzuständigkeiten angegangen und verfolgen jeweils eigenständige Zielrichtungen. Aus Sicht des RfII macht es allerdings Sinn, sie als idealerweise ineinandergreifende Zahnräder in einem Innovationen ermöglichenden umfassenden Datenökosystem zu begreifen. Beide Vorhaben gilt es wiederum mit vielen weiteren bereits vorhandenen Bausteinen im Bereich der nationalen, europäischen und internationalen Dateninfrastrukturen zu verzahnen und Wechselwirkungen zu berücksichtigen.

Einleitend sei darauf hingewiesen, dass heute in vielen Disziplinen und Forschungsfeldern Daten in internationaler Kooperation erzeugt oder angefordert werden. Deren Nutzung, Verwendung und auch Verwertung lässt sich nicht ohne weiteres allein in einem nationalen legislativen Rahmen regulieren. Zugriffe auf Forschungsdaten erfolgen zudem von überall her, also auch aus Staaten, in denen die Forschungsfreiheit und die Autonomie der Wissenschaft nicht in dem gleichen Maße gesichert sind, wie in den G7-Staaten und weiteren westlichen Ländern. Forschungsdaten können darüber hinaus in den Bereich der Dual Use-Güter fallen, die im Zweifel auch strengen Exportkontrollregeln unterliegen. Die Vielzahl der Rahmenbedingungen zeigt, dass nationale Regulierungen, die das Ziel verfolgen, mehr Anreize für das Datenteilen zu ermöglichen und Zugänglichkeit zu optimieren, in vielschichtiger Weise herausfordernd sind.

Des Weiteren dürfen die Vorhaben Forschungsdatengesetz und Dateninstitut in Deutschland und Europa nicht zu Doppelstrukturen oder unscharfen Verantwortlichkeiten führen. Derzeit bereits funktionierende freiwillige Arrangements im Datenaustausch – gerade innerhalb der Wissenschaft sowie zwischen dieser und außerwissenschaftlichen Partnern, etwa aus der Industrie – sollten nicht durch zusätzliche legislative Hürden erschwert werden. Der Status quo darf sich also nicht durch eine weitere Verrechtlichung existierender Vertrauensverhältnisse unabsichtlich verschlechtern. Hierbei gilt es auch zu berücksichtigen, dass ein nationales Forschungsdatengesetz hinsichtlich der **Harmonisierung mit parallelen Gesetzesvorhaben** auf der EU-Ebene (z. B. Data Act) und auf der Bundesebene (z. B. Gesundheitsdatennutzungsgesetz) hohen Anforderungen genügen muss. Das wirft die Frage auf, in welchem Zeitverhältnis die Vorlage eines Forschungsdatengesetzes einerseits zu den gegenwärtigen europäischen Aushandlungsprozessen zum Data Act und der legislativen Rahmung der entstehenden europäischen Datenräume und andererseits zur Errichtung eines Dateninstituts in Deutschland steht.¹ Letzteres soll laut Konzept vom 11.05.2023 ebenfalls zur Harmonisierung des gesellschaftlichen Datenteilens beitragen.² Die hierfür noch zu spezifizierenden Aufgabenbereiche und eine hierzu passende Governance werden derzeit in einem offenen Marktdialog erkundet, was auf eine längere Explorationsphase hindeutet. Zudem sollen Verbesserungen des Datenteilens in Use Cases und damit quasi aus einer praxisnahen „Bottom up“-Perspektive ermit-

1 Beispielsweise hat die EU-Kommission 2022 einen Vorschlag für eine Verordnung zum geplanten European Health Data Space (EHDS) auf den Weg gebracht. Weitere sektorspezifische EU-Rechtsetzungsvorhaben sind in diesem Zusammenhang zu erwarten.

2 Das vom BMWK und BMI in Abstimmung mit dem BMBF vorgelegte Konzept zum Aufbau des Dateninstituts ist abrufbar unter: bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/themen/it-digitalpolitik/dateninstitut/konzeptpapier_dateninstitut.pdf?__blob=publicationFile&v=3 (zuletzt geprüft am 23.06.2023). Erste Empfehlungen hierzu hatte die Gründungskommission für das Dateninstitut im Dezember 2022 formuliert. Siehe BMWK (2022) – Der Weg zu einem Dateninstitut für Deutschland. Zwischenbericht – Erste Empfehlungen der Gründungskommission, bmi.bund.de/SharedDocs/pressemitteilungen/DE/2022/12/dateninstitut.html (zuletzt geprüft am 23.06.2023).

telt werden, während ein Forschungsdatengesetz notwendigerweise „Top down“-Vorgaben macht – gegebenenfalls zu einem Zeitpunkt, zu dem die Auswertung der Use Cases noch nicht vorliegt.

Eine entscheidende Ausgangsfrage stellt sich in diesem Zusammenhang auch mit Blick auf die wesentliche Funktion eines Forschungsdatengesetzes: Soll es eine Klammer bilden, die auch Änderungen von anderen gesetzlichen Bestimmungen einschließt, die beispielsweise Zugänge der Forschung zu Daten der öffentlichen Hand aktuell behindern? Oder soll das Gesetz eine Lücke füllen, beispielsweise im bislang wenig regulierten Bereich des forschenden Zugangs zu privatwirtschaftlich und zivilgesellschaftlich gehaltenen Daten und umgekehrt, dem Zugang wirtschaftlicher und zivilgesellschaftlicher Akteure zu Forschungsdaten?

Hierzu wäre im Gesetzestext zunächst zu **klären, was unter Forschungsdaten genau verstanden werden soll**. Die bisherige Definition im Datennutzungsgesetz müsste in diesem Zusammenhang noch stärker operationalisiert werden, um z. B. in den komplexen Zugangsregimes an der Schnittstelle von Wissenschaft und Wirtschaft Rechtssicherheit zu schaffen.

Laut § 3, Abs. 10, Datennutzungsgesetz „sind Forschungsdaten Aufzeichnungen in digitaler Form, bei denen es sich nicht um wissenschaftliche Veröffentlichungen handelt und die im Laufe von wissenschaftlichen Forschungstätigkeiten erfasst oder erzeugt und als Nachweise im Rahmen des Forschungsprozesses verwendet werden oder die in der Forschungsgemeinschaft allgemein für die Validierung von Forschungsfeststellungen und -ergebnissen als notwendig erachtet werden“.

Die seitens des Rfll vorgeschlagene Begriffsklärung fasst unter Forschungsdaten auch (analoge wie digitale) Daten, „die den Entstehungsprozess der Primärdaten dokumentieren und kontextualisieren“. Diese so genannten „Sekundärdaten“ können „selbst wieder zu Primärdaten werden, was für den Lebenszyklus von Forschungsdaten von Bedeutung ist“.³ **Der Begriff Forschungsdaten schließt** darüber hinaus auch die sogenannten **Metadaten ein**, also alle Daten, die auf Prozesse der Qualitätssicherung (Kuratierung) sowohl im Forschungsprozess als auch in wissenschaftlichen Datenbanken und Repositorien zurückgehen und die Verwendbarkeit der Daten sicherstellen, beispielsweise deren Find- und Verknüpfbarkeit erleichtern.

Eine weitere notwendige Klärung – auch im Sinne einer wichtigen Randbedingung für ein Forschungsdatengesetz des deutschen Gesetzgebers – betrifft die Frage, welchen Regu-

3 Siehe Rfll (2016) – Leistung aus Vielfalt. Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland, Anhang A, 2.13, S. A 13.

lierungsspielraum auf nationaler Ebene der in Aushandlung befindliche europäische Data Act für den Bereich der Forschungsdaten und des Datenzugangs offen lassen wird.⁴

Ähnliches gilt auch für hiermit zusammenhängende, jedenfalls seitens verschiedener außerwissenschaftlicher Akteure im Vorfeld sowohl des Forschungsdatengesetzes als auch des Dateninstituts aufgeworfene Fragen nach wissenschaftlichen Datenbedarfen oder der Relevanz von Daten(-zugängen) für bestimmte wissenschaftliche Disziplinen, Felder oder Communities. Die Fragen „Welche Daten bzw. aus welchem Bereich braucht die Forschung Daten?“ und „Für wen sind Daten relevant?“ lassen sich nicht abschließend beantworten. Einerseits ist anzuerkennen, dass Daten aus unterschiedlichen gesellschaftlichen Teilbereichen immer auch Gegenstand sektorspezifischer Regulierungen sind (Gesundheit, Mobilität, Industrie etc.). Eine Fixierung solcher Sektoren, aus denen Daten für die öffentlich getragene Forschung von Belang sein sollen, ist in einem umfassenden Forschungsdatengesetz aber nicht hilfreich. **Regulatorische Vorgaben, die bestimmte Daten oder Datensektoren herausstellen, können aus wissenschaftlicher Perspektive schnell dysfunktional wirken oder im besten Falle immer nur vorläufig sein.** Gerade innovative Antworten auf Forschungsfragen beruhen oft auf Daten sehr heterogener Herkunft. Aus der Sicht des RfII lassen sich die Datenbedarfe der Wissenschaft in Deutschland daher nicht in sinnvoller Weise auf bestimmte, funktional beschriebene Datensorten begrenzen. Dies gilt in besonderem Maße auch deshalb, weil sich sowohl Forschungsfragestellungen als auch wirtschaftliche und gesellschaftliche Herausforderungen, zu denen die Wissenschaft Lösungskonzepte entwickelt, permanent weiterentwickeln. Um hier handlungsfähig zu bleiben, müssen die erforderlichen Daten Grundlagen immer wieder neu bewertet und aktualisiert werden können. Ein ebenso zukunfts offenes wie domänenübergreifendes Forschungsdatengesetz sollte insofern der Versuchung widerstehen, den aktuellen Status quo an Datenbedarfen oder Datensorten legislativ zu fixieren. Auch **Bedarfe nach Datenzugang und Anforderungen an die Datenqualität können sich rasch ändern oder neu entstehen.** Dies ist zum einen abhängig von technologischen und methodischen Innovationen und kann zum anderen – je nach wissenschaftlicher Disziplin und Domäne – in unterschiedlicher Geschwindigkeit ablaufen.

POSITIONEN DES RFII ZUR VERBESSERUNG DES DATENTEILENS AUS WISSENSCHAFTLICHER PERSPEKTIVE

Bereits vor den aktuellen Diskussionen um ein Forschungsdatengesetz und die Errichtung eines Dateninstituts hat der RfII Empfehlungen gegeben, die für beide Vorhaben von hoher Relevanz sind. Sie sind einerseits aus der Beschäftigung des Rates mit inhaltli-

4 Die Trilogverhandlungen zum Data Act wurden am 27. Juni 2023 unter schwedischer Ratspräsidentschaft erfolgreich abgeschlossen. Zum Zeitpunkt der Veröffentlichung dieses Diskussionsimpulses ist der Kompromissentwurf allerdings noch nicht öffentlich zugänglich.

chen Fragen der wissenschaftsfreundlichen Ausgestaltung von Zugangsmöglichkeiten zu Daten sowie deren Nutzung und Verwertung entstanden. Andererseits stammen sie aus der konkreten Beteiligung an Konsultationen zu europäischen Gesetzesvorhaben für den digitalen Binnenmarkt – insbesondere zum Data Governance Act (DGA) und zum Data Act (DA). Sie lassen sich ohne Substanzverlust auf Anforderungen an ein Forschungsdatengesetz und mittelbar auf die aus wissenschaftlicher Perspektive wünschenswerte Schwerpunktsetzung eines Dateninstituts übertragen.

In seiner Stellungnahme zum Data Governance Act (DGA, 2021) hat der RfII der EU mit Nachdruck empfohlen, für den breiten Geltungsbereich des DGA passende Zugangsregimes, etwa **Forschungs- bzw. Wissenschaftsklauseln**, zu etablieren, die es in einigen anderen Gesetzen (wie z. B. dem Netzwerkdurchsetzungsgesetz, NetzDG) bereits gibt, um auf diese Weise Wissenschaftlerinnen und Wissenschaftlern zum Zweck der Durchführung von Forschungsprojekten „ein unabdingbares Zugangsrecht zu den vorgehaltenen Daten zu gewähren“.⁵ Dabei hat der RfII mit Blick auf die durch den DGA angeregten, neuen Dateninfrastrukturen darauf hingewiesen, dass erstens „die Zwecke der Datenerhebung und Datenverwendung in Wirtschaft und Gesellschaft sehr differenziert sind“, so dass eine **One Size Fits All-Lösung nicht zielführend** wäre; zweitens **Haftungsfragen** einer besonderen Beachtung bedürfen sowie drittens **Akteure aus der Wissenschaft in Governancestrukturen** neuer Dateninfrastrukturen angemessen repräsentiert sein sollten.⁶ Diese Empfehlungen korrespondieren weitgehend mit dem Ansatz der Gründungskommission und dem Konzept zum Dateninstitut, entsprechende Lösungen – z. B. für intermediäre Datenmittler wie Datentreuhandstellen – aus Use Cases heraus zu entwickeln. Allerdings würde dies voraussetzen, dass Wissenschaftlerinnen und Wissenschaftler in den Konsortien hinreichend vertreten sind, die diese Use Cases bearbeiten und auswerten. Der RfII empfiehlt, bei der Auswahl der Konsortien, die die Aufgabenfelder für ein Dateninstitut entwickeln, auf eine angemessene Repräsentanz von Forschungsakteuren zu achten. Diese dringende Empfehlung hat auch den Hintergrund, dass es in der Wissenschaft bereits zahlreiche etablierte Einrichtungen gibt (z. B. in der NFDI, unter den RatSWD-akkreditierten Forschungsdatenzentren, im HPC-Bereich durch das NHR), die langjährige Erfahrungen mit der Ausgestaltung des vertrauensvollen Zugangs Dritter zu Forschungsdaten und auch umgekehrt des Zugangs zu Gesundheits-, Verwaltungs- und Wirtschaftsdaten für Forschungszwecke haben. Ein übergreifendes Dateninstitut sollte sich diese Erfahrungen zunutze machen, ohne die bereits existierenden Strukturen zu doppeln.

5 RfII (2021) – Stellungnahme zum Vorschlag eines Data Governance Acts (DGA) durch die EU-Kommission, S. 2.

6 Gegenüber der Empfehlung des RfII zur Einführung einer Forschungsklausel formuliert der DGA lediglich, dass u. a. für Forschungszwecke Daten öffentlicher Stellen zu ermäßigten Gebühren bzw. kostenlos zur Verfügung gestellt werden können (Art. 6, Abs. 4). Die vom RfII angeregten erweiterten Maßnahmen zur Qualitätssicherung in Bezug auf die Datenvermittlungsdienste, wie die Einführung eines Zertifizierungsverfahrens zumindest für ein hochqualitatives Teilssegment der Dienste, wurden im finalen Rechtstext in dieser Form nicht aufgegriffen.

Als Grundlage für ein erfolgreiches Schnittstellenmanagement über Infrastrukturgrenzen und sektorale Datensilos hinweg, regt der Rfll überdies nach wie vor eine **weitgehende Unabhängigkeit des Dateninstituts** gegenüber Bundesministerien und staatlichen Regulierungsbehörden an. Insgesamt muss sich ein Dateninstitut in eine Institutionenlandschaft einfügen, die schon heute in der digitalen Daten- und Informationsinfrastrukturpolitik durch eine Vielzahl von Expertenkommissionen und Beratungsgremien geprägt ist. Die Länderanalysen des Rfll haben gezeigt, dass diese Vielstimmigkeit kein Garant für eine kohärente Digitalpolitik zum Aufbau eines funktionierenden Datenökosystems ist. Deshalb wiederholt der Rfll seine 2022 im Positionspapier DATENPOLITIK, OPEN SCIENCE UND DATENINFRASTRUKTUREN: AKTUELLE ENTWICKLUNGEN IM EUROPÄISCHEN RAUM formulierte Empfehlung, dass „ein künftiges Dateninstitut mit nationalen Koordinierungsaufgaben [...] **zu einer schon heute unübersichtlichen Beratungslandschaft in Deutschland nicht einfach dazu treten, sondern die verschiedenen Stränge zu einem komplementären Paket von Beratungsangeboten bündeln**“ und dazu beitragen sollte, verschiedene Sektoren (Staat, Wirtschaft, Gesellschaft und Wissenschaft) und Domänen miteinander zu verbinden.⁷ Zur Stärkung des in beiden Gründungsdokumenten empfohlenen Bottom-up-Ansatzes betrachtet der Rfll die Entwicklung von niederschweligen, agilen und vor allem markt-basierten Instrumenten für die Förderung des sektorübergreifenden Datenteilens als hilfreich. Dazu könnte z.B. die Förderung von Protagonisten (ähnlich dem Sovereign Tech Fund), die Vergabe von Innovations-Vouchern oder Stipendien zur Bewältigung konkreter Herausforderungen datenbasierter Anwendungen gehören.

Zu den Aufgaben, die ein mögliches Dateninstitut im Interesse der Forschung prioritär angehen sollte, zählt der Rfll die **Beseitigung von Hindernissen bezüglich des Zugangs und im Umgang mit Daten**. Diese betreffen in Deutschland u.a. die mangelnde Auffindbarkeit von Daten, die fehlende Rechtssicherheit bei der Verknüpfung von Daten, die Abgrenzungsproblematik personenbezogener von nicht-personenbezogenen Daten, nicht rechtssichere Leitlinien zur Anonymisierung sowie fehlende standardisierte Datenzugänge u.a. zu privatwirtschaftlichen Daten.

Allgemein ist die Rechtsunsicherheit bzw. ein unklarer Rechtsstatus von Daten in Deutschland ein Hindernis für deren sektorenübergreifende Nutzung – in der Forschung selbst, aber noch mehr mit Blick auf Daten, die von anderen Akteuren generiert werden. Für die Nachnutzung von Daten in der Forschung ist ein Verzicht auf Lizenzierung nicht förderlich. Im Gegenteil: Eindeutige Lizenzierung gibt Datengebenden und Datennutzenden Sicherheit im Umgang mit Daten und Metadaten. Darüber hinaus ist die Frage, ab wann das Urheberrecht greift, noch nicht grundsätzlich geklärt. Wann ist die Schwelle einer geistigen Schöpfung erreicht – bei Kommentierung, Kontextualisierung, Auswahl, An-

7 Vgl. Rfll (2022) – Datenpolitik, Open Science und Dateninfrastrukturen: Neue Initiativen im europäischen Raum, S. 77.

reicherung mit Normdaten? Der Wissenschaft könnte durch rechtliche Klärungsschritte erleichtert werden, Datenprodukte zu entwickeln und diese sektorenübergreifend bereitzustellen.

Der RfII hat schon früher dafür plädiert, der **Wissenschaft unter bestimmten Voraussetzungen einen eigenen Zugriff auf privatwirtschaftliche Daten einzuräumen**.⁸ Dies liegt im öffentlichen Interesse. Soll die Wissenschaft zur Lösung großer gesellschaftlicher Herausforderungen beitragen können, darf man sie nicht ausschließlich auf die Arbeit mit selbst erhobenen Daten reduzieren. Die öffentlich getragene Wissenschaft würde sonst in relativ kurzer Zeit hinter das zurückfallen, was sie in Bereichen datenintensiver Forschung (und etwa auch im Bereich der Künstlichen Intelligenz) leisten kann und was man zu Recht von ihr erwartet. Im Zuge einer Regulierung, die der Wissenschaft Datenzugang ermöglicht, könnte aber durchaus festgelegt werden, dass über die wissenschaftliche Datenverwendung auf Nachfrage Rechenschaft abzulegen ist. Als einen weiteren wichtigen Punkt sieht der RfII die Begrenzung eventueller Kosten, die der Forschung für den Datenzugang entstehen: Die öffentlich finanzierte Wissenschaft sollte privatwirtschaftlichen Datengebern höchstens eine Aufwandsentschädigung für die Datenbereitstellung erstatten müssen. Darüber hinaus gehende Kosten würden auf Forschungsvorhaben prohibitiv wirken und letztlich den Steuerzahler belasten. Gleichwohl erkennt der RfII an, dass berechnete Geschäftsgeheimnisse und auf Daten basierende Geschäftsmodelle im Unternehmensbereich nicht verletzt bzw. konterkariert werden dürfen. Hier empfiehlt der RfII mit Blick auf ein Forschungsdatengesetz gut abzuwägen, inwieweit bzw. welche gesetzlichen Vorgaben hier für die Wissenschaft tatsächlich Vorteile des Datenzugangs bringen. Bisherige **freiwillige Kooperationen zwischen Wissenschaft und Wirtschaft sollten nicht restriktiver gestaltet werden**, denn dann werden sie insbesondere für die Wirtschaft unattraktiver.

In seinem Positionspapier NUTZUNG UND VERWERTUNG VON DATEN IM WISSENSCHAFTLICHEN RAUM (2021) hat der RfII nachdrücklich empfohlen, **Datentreuhandmodelle zu fördern und wissenschaftsfreundlich auszugestalten** sowie Verfahren der Qualitätssicherung aus der Wissenschaft auch außerhalb des Wissenschaftssystems maßvoll zu adaptieren. Konkret wären u. a. Forschungsklauseln gesetzlich zu verankern, Forschungsdatenzentren auszubauen sowie Zertifizierungs- bzw. Akkreditierungsniveaus anhand der Güte- und Schutzbestimmungen der Daten zu bestimmen.⁹ In diesem Zusammenhang stellt sich die Frage, ob ein nationales Forschungsdatengesetz eine Verbesserung im Vergleich zu einem Pluriversum an Forschungsklauseln in anderen Gesetzen sein kann. **Übergeneralisierte Lösungen hat der RfII bislang mit Verweis auf die differenzierten Bedarfe**

8 Vgl. RfII (2022) – Stellungnahme zum Vorschlag der EU-Kommission für eine „Verordnung über harmonisierte Vorschriften für den fairen Zugang zu Daten und deren Verwendung“ (Data Act).

9 Vgl. RfII (2021) – Nutzung und Verwertung von Daten im wissenschaftlichen Raum. Empfehlungen zur Ausgestaltung von Datendiensten an der Schnittstelle zwischen Wissenschaft und Wirtschaft, Empfehlung 5.3, S. 77f.

wissenschaftlicher Disziplinen, Forschungsfelder und Communities abgelehnt. Ob bzw. wie ein Forschungsdatengesetz sowohl diese differenzierten Bedarfe abbilden kann als auch die unterschiedlichen Wettbewerbsbedingungen in ein allen Seiten genügendes Verhältnis setzen kann, ist eine wichtige Frage. Gleiches gilt für die Berücksichtigung der Datenschutzbedürfnisse von datenerhebenden und -verarbeitenden Unternehmen, aber auch von datengebenden Privatpersonen. Hinzu kommt, dass wichtige Bereiche der wissenschaftlich gewünschten erleichterten Datenzugänge oder Datenverknüpfungsmöglichkeiten in sehr unterschiedlichen Gesetzen (z.B. Bundesstatistikgesetz, Steuergesetze, Urheberrechtsgesetz) und auf unterschiedlichen politischen Ebenen (z.B. das Datenschutzrecht in Europa, auf Bundes- und auf Länderebene) bereits geregelt sind oder zurzeit geregelt werden (z.B. Gesundheitsdatennutzungsgesetz). Auch die unterschiedliche Gestaltung von „Soft Law“ durch eine nicht harmonisierte Praxis der Datenschutzbehörden der Länder erschwert die Forschung mit öffentlich erhobenen Daten in Deutschland. Wirklich wirksam kann ein Forschungsdatengesetz nur werden, wenn es alle diese Rahmenbedingungen (mit) adressiert.

Der RfII empfiehlt daher, zunächst zu prüfen, welche bisherigen Regulierungen durch andere Gesetze des Bundes und der Länder einem wissenschaftlichen Zugang zu Daten bzw. der Verknüpfung von Daten, die unter verschiedene Normbereiche fallen, im Wege stehen bzw. welche Gesetze Vorgaben enthalten, die durch Forschende berücksichtigt werden müssen. Bestehende Restriktionen z.B. im Statistikgesetz des Bundes lassen sich voraussichtlich nicht durch ein neues Forschungsdatengesetz „überschreiben“. Dies gilt auch für Vorgaben des deutschen Außenwirtschaftsrechts, die sich potenziell auf das Teilen von Daten prohibitiv auswirken können. Entsprechend könnte auf Grundlage einer entsprechenden Evaluierung anderer Gesetze mit Vorgaben zu Datenzugängen geprüft werden, ob ein **Forschungsdatengesetz als „Artikelgesetz“** zu formulieren wäre. Dieses könnte Forschungsklauseln in bereits bestehende Gesetze eintragen bzw. die Verknüpfung von Datenbeständen, die unterschiedlichen Rechtskreisen zugeordnet sind, zu Forschungszwecken erleichtern. Entsprechende **Bemühungen, eine Harmonisierung auch mit den Landesgesetzgebern zu erreichen, müssten folgen.**

Der RfII bekräftigt, dass ein Forschungszugang zu öffentlich erhobenen und gehaltenen Daten durch die Überwindung bisheriger gesetzlicher Restriktionen und im Einklang mit einem forschungsfreundlich verstandenen Datenschutzrecht auch im Gemeinwohlinteresse unabdingbar ist. Insbesondere für die seitens Politik und Öffentlichkeit zu Recht geforderte evidenzbasierte Politikberatung werden Datenzugänge für eine Forschung benötigt, die nur in der öffentlich geförderten Wissenschaft in hinreichend übergreifender Form, neutral sowie in gesamtgesellschaftlicher Perspektive erfolgen kann. Für Kooperationen zwischen Wissenschaft und Wirtschaft präferiert der RfII **freiwillige Vereinbarungen – insbesondere im Bereich der Kleinen und Mittleren Unternehmen.** Für global agierende Großkonzerne der digitalen Wirtschaft können im Einklang mit dem europäischen *Digital Markets Act* und *Digital Services Act* auch rechtliche Ansprüche auf Datenzugang

der öffentlich getragenen Wissenschaft zu Daten von „Hyperscalern“ festgeschrieben werden. Umgekehrt muss allerdings zwingend vermieden werden, dass qua Rechtsetzung oder -auslegung Einfallstore eröffnet werden, mit denen durch öffentlich getragene Forschung generierte Daten ohne die Zustimmung ihrer Urheber transferiert und monetarisiert werden.¹⁰ In diesem Sinne **sollte ein Forschungsdatengesetz dezidiert als ein „Datenzugangsgesetz für die Forschung“¹¹ konzipiert sein.** Der Rfll unterstützt diese Forderung der DFG ausdrücklich.

POSITIONEN DES RFLL ZUM ZUGANG ZU DATEN AUS DER ÖFFENTLICH GETRAGENEN FORSCHUNG ZUM ZWECKE DER NUTZUNG UND VERWERTUNG IN ANDEREN GESELLSCHAFTLICHEN SEKTOREN

Der Rfll bekennt sich grundsätzlich zum Paradigma eines offenen Zugangs zu Daten aus der Wissenschaft (Open Science). Schon 2016 hat er allerdings darauf hingewiesen, dass „der Umgang mit der Systemgrenze zwischen Wissenschaft und Wirtschaft eine Reihe von Fragen aufwirft. Gerade um unter den Akteuren innerhalb der Wissenschaft eine möglichst offene Kultur des Teilens von Forschungsdaten zu etablieren, müssen rechtliche Schranken errichtet werden, die auch global verhindern, dass Akteure einen in der Wissenschaft niedrigschwiligen Zugang zu Daten nutzen, um Eigentum an dem zu begründen, was frei (im Sinne eines wissenschaftlichen Gemeinschaftsgutes) unter Forschenden zirkulieren soll.“¹²

Deutlicher hat der Rfll diesen Gedanken des rechtlichen Schutzes von Offenheit vor einer kommerziellen Unterminierung durch ungewollte Aneignung und Verwertung 2019 in seiner Stellungnahme zu aktuellen Entwicklungen rund um Open Data formuliert. Hier gibt der Rat zu bedenken, dass die Strategie „für den freien Zugang zu Forschungsdaten auf wissenschaftsadäquate Modelle des Teilens von Daten setzen und anstelle eines quantitativen Wachstums vor allem das qualitative Ziel hochwertiger Datenbestände und Datendienste verfolgen“ muss. **Eine „undifferenzierte ‚Publizitätspflicht‘ – etwa auch für alle vorläufigen Zwischenprodukte auf dem Weg zu einem Ergebnis“ lehnt der Rat ab,** da dies „dem spezifischen Leistungsanspruch und der Qualitätsverantwortung der Wissenschaft nicht“ entspricht.¹³ Dem Rfll ist es wichtig zu betonen, dass es sich bei Datenpro-

10 Vgl. Rfll (2021) – Nutzung und Verwertung von Daten im wissenschaftlichen Raum. Empfehlungen zur Ausgestaltung von Datendiensten an der Schnittstelle zwischen Wissenschaft und Wirtschaft, S. 79.

11 DFG (2023) – Die Wissenschaft braucht ein Forschungsdatengesetz, www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/forschungsdaten/stellungnahme_forschungsdatengesetz.pdf (zuletzt geprüft am 01.06.2023), S. 1 (Hervorhebung durch den Rfll).

12 Rfll (2016) – Leistung aus Vielfalt. Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland, S. 35.

13 Rfll (2019) – Stellungnahme zu aktuellen Entwicklungen rund um Open Data und Open Access, S. 2, vgl. auch Rfll (2019) – Herausforderung Datenqualität. Empfehlungen zur Zukunftsfähigkeit von Forschung im digitalen Wandel, Empfehlung 4.4.2, S. 90.

dukten, welche die Wissenschaft öffentlichen Stellen, der Wirtschaft oder der Zivilgesellschaft zur Verfügung stellt, um qualitätsgesicherte, verlässliche und in diesem Sinne bereits als Ergebnis von Forschung zu klassifizierende Daten handeln muss.

Im Positionspapier NUTZUNG UND VERWERTUNG VON DATEN IM WISSENSCHAFTLICHEN RAUM (2021) hat der RfII zwischen dem Zugang zu Daten, ihrer wissenschaftlichen Nutzung und erst im dritten Schritt ihrer Verwertung nach Weitergabe an die Wirtschaft oder an eine auf kommerziellen Arrangements basierenden Öffentlichkeit unterschieden. Auf dieser dritten Stufe der Verwertung von Daten greifen Regelungen, die nicht mehr dem wissenschaftlichen *Common Sense* (als wesentlichem Bestandteil einer guten wissenschaftlichen Praxis) unterliegen. Hier kommen rechtliche Regelungen in Form „formalisierter Vereinbarungen ins Spiel (Lizenzen, Verträge, Vereinbarungen etc.).“¹⁴ Von diesen Feststellungen ausgehend hat der RfII u.a. Empfehlungen zur staatlichen Regulierung des Marktes für kommerzielle und öffentlich finanzierte Datendienste ausgesprochen.

Zunächst empfiehlt der RfII:¹⁵

- Frei- und Schutzräume für innerwissenschaftliche Nutzung und komplikationsloses Teilen von Daten zu definieren;
- Verwendung von Forschungsdaten erst nach Einwilligung der datenerzeugenden Akteure zu gestatten;
- fragile Voraussetzungen der Autonomie von Forschung in der Gesetzgebung zu berücksichtigen.

In diesem Zusammenhang bekräftigt der Rat „den **Grundsatz, dass mit öffentlichen Mitteln erzeugte wissenschaftliche Daten auch öffentlich zugänglich bleiben müssen**“. Er begrüßt ausdrücklich das Teilen von Daten mit außerwissenschaftlichen Interessenten, weist allerdings darauf hin, dass „das Interesse Dritter am Zugang zu Forschungsdaten im Forschungsprozess bzw. vor Abschluss einer jeweiligen wissenschaftlichen Studie auch weiterhin die Regeln und Regulierungen von wissenschaftlichen Fachgemeinschaften und Professionszusammenschlüssen zu respektieren hat.“¹⁶ Mit Blick auf die Ausgestaltung eines Forschungsdatengesetzes heißt dies: „**Frei- und Schutzräume für innerwissenschaftliches Datenteilen müssen im Rahmen aktueller und künftiger datenbezogener Rechtsetzungsprozesse [...] klar formuliert sein. Eine diesbezüglich zu fragmentierte und uneindeutige Rechtslage würde auf die Bereitschaft, Daten zu veröffentlichen, hemmend wirken.**“¹⁷

14 RfII (2021) – Nutzung und Verwertung von Daten im wissenschaftlichen Raum. Empfehlungen zur Ausgestaltung von Datendiensten an der Schnittstelle zwischen Wissenschaft und Wirtschaft, S. 12, 16.

15 Siehe ebd., Empfehlung 5.1, S. 74ff.

16 Ebd., S. 75.

17 Ebd., S. 76.

Ein unkonditionierter und pauschal geregelter Zugang außerwissenschaftlicher Akteure zu Forschungsdaten zu wenig spezifizierten Zwecken ist aus Sicht des RfII nicht sinnvoll. Vielmehr sollte die Wissenschaft ihre Daten in Form von Datenprodukten zur Verfügung stellen können, die eine gute und qualitätswahrende außerwissenschaftliche Nutzung dann auch tatsächlich sichern. Es kommt in diesem Zusammenhang auf konkrete Fragestellungen an, aus denen sich etwa ableiten lassen müsste, welchen Grad der Verarbeitung (z.B. Anonymisierung, Pseudonymisierung, Kuratierung allgemein) Forschungsdaten haben müssen, um für außerwissenschaftliche Zwecke überhaupt brauchbar zu sein, mit welchem Aufwand bzw. mit welchen Kosten dies verbunden wäre und ob die Zwecke selbst mit wissenschaftlichen Wertvorstellungen und guter wissenschaftlicher Praxis vereinbar sind.

Bei der rechtlichen Ausgestaltung der Kooperationsbeziehungen zwischen innerwissenschaftlichen Akteuren und Externen hält der RfII für den Fall der Übergabe von Forschungsdaten an vermarktende Partner die vertragliche Sicherung einer kostenfreien wissenschaftlichen (Weiter-)Nutzung der Daten für unabdingbar. Werden Forschungsdaten durch Dritte verwertet bzw. vermarktet, **darf der Wissenschaft ein dauerhafter und vollständiger Zugriff auf die von ihr erzeugten Daten nicht entzogen werden**. Auch wenn die kommerziellen Kooperationspartner die Daten anreichern oder auf andere Weise veredeln, sollen gegenüber den ursprünglichen wissenschaftlichen Datengebern lediglich maßvolle Gebühren zulässig sein.¹⁸ Sofern in einem Forschungsdatengesetz die Einschaltung von neutralen Datentreuhändern beispielsweise beim Transfer privatwirtschaftlicher Daten an Forschungsakteure vorgesehen werden sollte, empfiehlt der RfII, hinsichtlich der Kosten einer solchen Vermittlungsleistung eine wissenschaftsfreundliche Regelung zu finden, die sich beispielsweise auch in die Projektförderung einbauen ließe. Hohe Gebühren bei der Indienstnahme eines Treuhänders würde die entsprechenden Forschungsvorhaben ansonsten erschweren.

Dies verweist auch auf die Aufgaben eines möglichen Dateninstituts. Es gilt für Datenintermediäre nicht nur eine rechtliche Grundlage, sondern auch geeignete, gemeinwohlorientierte Geschäftsmodelle zu entwickeln. Von daher sollte ein künftiges Dateninstitut neben technischer und rechtlicher Expertise auch Know How in der Frage guter Betriebsmodelle für die Dateninfrastrukturen bereithalten, welche die öffentlich getragene Wissenschaft an ihren Schnittstellen zu anderen gesellschaftlichen Bereichen so dringend benötigt.

18 Vgl. Empfehlung 5.4, ebd., S. 79, vgl. auch RfII (2019) – Herausforderung Datenqualität. Empfehlungen zur Zukunftsfähigkeit von Forschung im digitalen Wandel, Empfehlung 4.4.4, S. 91.

LITERATURVERZEICHNIS

BMWK – Bundesministerium für Wirtschaft und Klimaschutz; BMI – Bundesministerium des Innern und für Heimat (2023): Konzept zum Aufbau des Dateninstituts, Berlin, bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/themen/it-digitalpolitik/dateninstitut/konzeptpapier_dateninstitut.pdf?__blob=publicationFile&v=3.

BMWK – Bundesministerium für Wirtschaft und Klimaschutz; Gründungskommission (2022): Der Weg zu einem Dateninstitut für Deutschland. Zwischenbericht – Erste Empfehlungen der Gründungskommission, Berlin, bmi.bund.de/SharedDocs/pressemitteilungen/DE/2022/12/dateninstitut.html (zuletzt geprüft am 23.06.2023).

DFG – Deutsche Forschungsgemeinschaft (2023): Die Wissenschaft braucht ein Forschungsdatengesetz, Bonn, dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/forschungsdaten/stellungnahme_forschungsdatengesetz.pdf.

RfII – Rat für Informationsinfrastrukturen (2022): Datenpolitik, Open Science und Dateninfrastrukturen: Neue Initiativen im europäischen Raum. Bericht und Empfehlungen, Göttingen, URN: [urn:nbn:de:101:1-2021090875](https://nbn-resolving.org/urn:nbn:de:101:1-2021090875).

RfII – Rat für Informationsinfrastrukturen (2022): Stellungnahme zum Vorschlag der EU-Kommission für eine „Verordnung über harmonisierte Vorschriften für den fairen Zugang zu Daten und deren Verwendung“ (Data Act), Göttingen, URN: [urn:nbn:de:101:1-2021090857](https://nbn-resolving.org/urn:nbn:de:101:1-2021090857).

RfII – Rat für Informationsinfrastrukturen (2021): Nutzung und Verwertung von Daten im wissenschaftlichen Raum. Empfehlungen zur Ausgestaltung von Datendiensten an der Schnittstelle zwischen Wissenschaft und Wirtschaft, Göttingen, URN: [urn:nbn:de:101:1-2020052673](https://nbn-resolving.org/urn:nbn:de:101:1-2020052673).

RfII – Rat für Informationsinfrastrukturen (2021): Stellungnahme zum Vorschlag eines Data Governance Acts (DGA) durch die EU-Kommission, Göttingen, URN: [urn:nbn:de:101:1-2020052654](https://nbn-resolving.org/urn:nbn:de:101:1-2020052654).

RfII – Rat für Informationsinfrastrukturen (2019): Herausforderung Datenqualität – Empfehlungen zur Zukunftsfähigkeit von Forschung im digitalen Wandel, 2. Aufl., Göttingen, URN: [urn:nbn:de:101:1-2019112011541657732737](https://nbn-resolving.org/urn:nbn:de:101:1-2019112011541657732737).

RfII – Rat für Informationsinfrastrukturen (2019): Stellungnahme zu den aktuellen Entwicklungen rund um Open Data und Open Access, Göttingen, URN: [urn:nbn:de:101:1-2019051612340151989161](https://nbn-resolving.org/urn:nbn:de:101:1-2019051612340151989161).

RfII – Rat für Informationsinfrastrukturen (2016): Leistung aus Vielfalt. Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland, Göttingen, URN: [urn:nbn:de:101:1-201606229098](https://nbn-resolving.org/urn:nbn:de:101:1-201606229098).

IMPRESSUM

Rat für Informationsinfrastrukturen (RfII)

Geschäftsstelle

Papendiek 16

37073 Göttingen

Tel. 0551-3927050

E-Mail info@rfii.de

Web www.rfii.de

ZITIERVORSCHLAG

Rat für Informationsinfrastrukturen: Diskussionsimpuls. Anforderungen an die Ausgestaltung eines Forschungsdatengesetzes und eines Dateninstituts, Göttingen 2023, 12 S.

Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung – keine Bearbeitung 4.0 Lizenz (CC BY-ND).



Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind über URN [urn:nbn:de:101:1-2023021729](https://nbn-resolving.org/urn:nbn:de:101:1-2023021729) abrufbar.