FORSCHUNGSDATENRÄUME[1]

I. Was ist ein Forschungsdatenraum?

Nach einer Definition eines Forschungsdatenraums lässt sich lange suchen. Dies liegt daran, dass es eigentlich nicht „den Forschungsdatenraum“ gibt. Letztlich handelt es sich bei Forschungsdatenräumen um Data Spaces, in denen ein Austausch von **Forschungsdaten** stattfindet. Dies ist durch Datenbanken, Zusammenschlüsse von Forschern, Staat und Industrie als auch **mit Datenzugangsökosystemen über Zugangsansprüche zu Daten** (S. 6-8) möglich. Den Forschungsdatenraum für die gesamte Wissenschaft gibt es (noch (S. 3)) nicht. Stattdessen gibt es **spezifische Forschungsdatenräume**, die im Rahmen einzelner Kooperationen einen Datenaustausch zwischen unterschiedlichen Institutionen ermöglichen. Jedoch arbeiten Projekte wie die NFDI (S. 2-3), die EOSC (S. 3-4) und die GOSC (S. 10) auf das Ziel einer einzigen Anlaufstelle für Forschungsdaten hin.

II. Was sind Forschungsdaten und Forschungsdatenmanagement?

Die Plattform forschungsdaten.info, die vom baden-württembergischen Begleit- und Weiterentwicklungsprojekt für Forschungsdatenmanagement (bw2FDM) betrieben wird, definiert Forschungsdaten als *„**(digitale) Daten, die während wissenschaftlicher Tätigkeit** (z.B. durch Messungen, Befragungen, Quellenarbeit) **entstehen**. Sie bilden eine Grundlage wissenschaftlicher Arbeit und dokumentieren deren Ergebnisse. Daraus ergibt sich ein disziplin- und projektspezifisches Verständnis von Forschungsdaten mit unterschiedlichen Anforderungen an die Aufbereitung, Verarbeitung und Verwaltung der Daten: dem sogenannten Forschungsdatenmanagement. Bisweilen wird auch zwischen Primärdaten und Metadaten unterschieden, wobei letztere je nach Fachbereich oft nicht als Forschungsdaten im engeren Sinne gelten.“*[2]

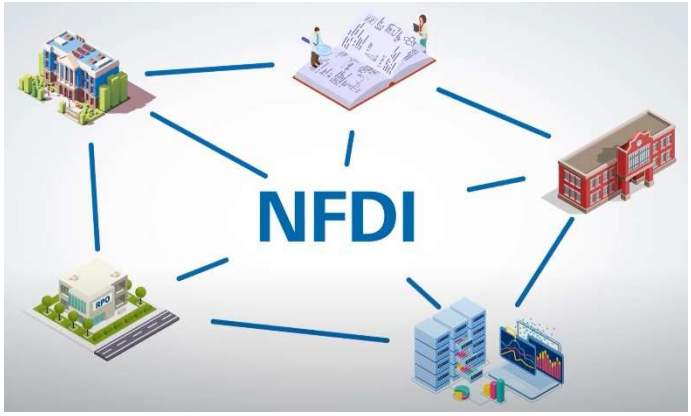
Unter **Forschungsdatenmanagement** ist die *„Transformation, Selektion und Speicherung von Forschungsdaten mit dem Ziel, diese langfristig und nachprüfbar zu halten“* zu verstehen.[3]

Forschungsdaten sollen also systematisch geordnet, langfristig gespeichert, offen verfügbar und national und international verknüpft werden (Abbildung).[4]

[1] Der Beitrag entstand im Rahmen von Recherchetätigkeiten des Autors, Nicolas Kutschera, im September und Oktober 2022 als wissenschaftlicher Mitarbeiter am Lehrstuhl für Strafrecht, Strafprozessrecht, Rechtstheorie, Informationsrecht und Rechtsinformatik bei Prof. Dr. Dr. Eric Hilgendorf. Der Text wurde im August 2023 überarbeitet. Der Autor ist für den Inhalt allein verantwortlich.
[2] Glossar der Seite forschungsdaten.info zum Begriff Forschungsdaten, https://forschungsdaten.info/praxis-kompakt/glossar/#c269821, genauso auch Prof. Dr. Peter Schirmbacher und Maxi Kindling: *Aust, Pamela/ Helbig, Kerstin/ Rosenbaum, Anja,* Was sind Forschungsdaten? Ein Online-Tutorial der Forschungsdatenmanagement-Initiative der Humboldt-Universität zu Berlin, 02.12.2016, https://rs.cms.hu-berlin.de/dataman/pages/view.php?ref=90, jeweils zuletzt abgerufen am 22.08.2023.
[3] Glossar der Seite forschungsdaten.info zum Begriff Forschungsdatenmanagement, https://forschungsdaten.info/praxis-kompakt/glossar/#c269836, zuletzt abgerufen am 22.08.2023.
[4] Abbildung ist ein Screenshot des Videos der DFG zur NFDI vom 07.10.2021, https://youtu.be/uJ01g9m8uE4, zuletzt abgerufen am 22.08.2023.

Forschungsdaten als wesentliche Grundlage des wissenschaftlichen Arbeitens müssen zur Qualitätssicherung, Nachvollziehbarkeit und Reproduzierbarkeit der eigentlichen Arbeit stets nachhaltig gesichert und bereitgestellt werden. Denn die Forschungsdaten können dadurch besser für andere Forschungsvorhaben weiterverwendet werden.[5] Etabliert ist das mittlerweile 30 Jahre alte Datenrepositorium PANGAEA, das nach einer Initiative einer Vielzahl von Wissenschaftlern vom Alfred-Wegener-Instituts betrieben wird und geologische und umweltbezogene Daten offen und frei bereithält.[6]

Ähnlich erfolgreich den Datenaustausch zu ermöglichen sind die Ziele der *Nationalen Forschungsdaten Infrastruktur* (NFDI). Sie möchte eine nachhaltige Struktur für das Forschungsdatenmanagement aufbauen. Ziel der NFDI ist es dabei auch ein Netzwerk zwischen Forschungseinrichtungen, Forschern, Datenbanken und Auftraggebern aufzubauen (Abbildung).[7]



III. Datenräume auf nationaler Ebene

Der vom RatSWD und der DFG geförderte Verein Nationale Forschungsdaten Infrastruktur (NFDI) besitzt zu 26 **„NFDI-Konsortien“**, die gemeinsam in ihren einzelnen Themenbereichen Daten austauschen.[8] Vom RatSWD wurden sogar **42 Forschungsdatenzentren akkreditiert**.[9] Die Weiterentwicklung der Nationalen Forschungsdateninfrastruktur und einen Europäischen Forschungsdatenraum wollte auch die Ampel ausweislich ihres **Koalitionsvertrag**es vorantreiben.[10]

Was ein **NFDI-Konsortium** ist und was es macht, soll am Beispiel des „Text+“-Konsortiums erklärt werden. Text+ möchte sprach- und text-basierte Forschungsdaten austauschen, um die Nutzungsfreundlichkeit der Daten und Werkzeuge, verbunden mit der rechtskonformen Nutzung von Sprach- und Textdaten, zu verbessern. Dabei soll eine ortsverteilte Infrastruktur für Sprach- und Textdaten erstellt, Forschende bei der Erstellung, Nachnutzung und Erhaltung von Sprach- und Textdaten unterstützt, innovative Forschung durch leichten Zugang zu Forschungsdaten und Werkzeugen ermöglicht werden und das unter enger Zusammenarbeit mit beteiligten Fachcommunities durch Portfolioerweiterung, Training oder Workshops. Die Aufgabe des Konsortiums ist Reference Implementation, Portfolio Development, Standardisation Activities, Community Activities und Software-Services. Sprecher des Konsortiums ist Prof. Dr. Erhard Hinrichs des Leibniz Institut für Deutsche Sprache Mannheim. Beteiligt sind eine Vielzahl von Universitäten, auch die JMU, einige Stiftungen als auch Akademien der Wissenschaften. Um

---

[5] *DFG,* Leitlinien zum Umgang mit Forschungsdaten, 30. September 2015.
[6] *Diepenbroek, Michael/ Grobe, Hannes,* PANGAEA - Network for Geological and Environmental Data. Pages.1999, https://doi.org/10.22498/pages.7.1.9.
[7] Abbildung ist ein Screenshot des Videos der DFG zur NFDI vom 07.10.2021, https://youtu.be/uJ01g9m8uE4, zuletzt abgerufen am 22.08.2023.
[8] Seite des NFDI, Konsortien, https://www.nfdi.de/konsortien/, zuletzt abgerufen am 22.08.2023.
[9] Sämtliche akkreditierte Forschungsdatenzentren finden sich unter: https://www.konsortswd.de/datenzentren/alle-datenzentren/, zuletzt abgerufen am 22.08.2023
[10] *SPD/ BÜNDNIS 90 / DIE GRÜNEN/ FDP,* Mehr Fortschritt wagen. Bündnis für Freiheit, Gerechtigkeit und Nachhaltigkeit. Koalitionsvertrag 2021 – 2025 zwischen der Sozialdemokratischen Partei Deutschlands (SPD), BÜNDNIS 90 / DIE GRÜNEN und den Freien Demokraten (FDP), s. 21.

seine Ziele umzusetzen hat Text+ mehrere Boards und Koordinierende Ausschüsse.[11] Die NFDI selbst konzentriert sich zunächst auf die technische und auch personelle Vernetzung entlang der einzelnen Disziplinen. Mit Hilfe eines NFDI-Direktorats sollen die einzelnen Initiativen auch ineinanderwirken.[12]

Die vom **RatSWD** akkreditierten **Forschungsdatenzentren** sind anders organisiert. Hier bietet beispielsweise das Forschungsdatenzentrum des Bundesamtes für Migration und Flüchtlinge Zugang zu Daten des Ausländerzentralregisters (AZR) an, einem Register, was eigentlich der Unterstützung der Behörden dient, die mit der Durchführung von ausländer- und asylrechtlichen Vorschriften betraut sind.[13] Das Forschungsdatenzentrum des Robert Koch-Instituts stellt eine Vielzahl an faktisch anonymisierten Gesundheitsdaten, Telefonischen Gesundheitssurveys, Kinder- und Jugendgesundheitssurveys, etc. bereit.[14]

Seit Mai 2021 läuft das vom BMBF geförderte Projekt **FAIR Data Spaces**, dass die föderierte, sichere Dateninfrastruktur von **Gaia-X** und **NFDI** zu einem **gemeinsamen cloud-basierten Datenraum für Industrie und Forschung** unter Einhaltung der FAIR-Prinzipien verbinden möchte. Nach diesen Prinzipien sollen die Daten *auffindbar, zugreifbar, interoperabel und wiederverwendbar* (**f**indable, **a**ccessible, **i**nteroperable, **r**eusable) sein. Diese Verknüpfung soll die Zusammenarbeit von Industrie und Forschung durch eine bessere technische Grundlage fördern.[15]

IV. Datenräume und Regulierung auf europäischer Ebene

Forschungsdatenräume auf europäischer Ebene sind die logische Folge der Zielbestrebung des Art. 3 Abs. 3 Satz 3 EUV, nachdem die EU *„den **wissenschaftlichen und technischen Fortschritt“*** **zu fördern** hat und den Art. 179 ff. AEUV, die diese Zielbestrebungen durch einzelne Maßnahmen ausfüllen. Insbesondere Art. 180 Buchstabe c AEUV, der die *„Verbreitung und Auswertung der Ergebnisse der Tätigkeiten auf dem Gebiet der Forschung der Union, technologischen Entwicklung und Demonstration“* benennt ist taugliche Ermächtigungsgrundlage der EU-Organe, um gesetzgeberisch tätig zu werden. Die **zentrale Rolle der Forschung** in der EU wird **auch im Datenschutzrecht**, insbesondere in Art. 89 DSGVO, **anerkannt**, bedeute aber keine „*zügellose Freiheit der Forschung*“.[16]

1. European Open Science Cloud (EOSC)

Das BMBF nennt in seinem *BMBF-Aktionsplan Forschungsdaten* vom Oktober 2020 den Start einer europäischen offenen Wissenschaftscloud, die die Wettbewerbsfähigkeit der deutschen und europäischen Wirtschaft durch Schaffung einer **zuverlässigen und offenen Umgebung für die gemeinsame Speicherung, Nutzung und Wiederverwendung von Daten durch die wissenschaftliche Gemeinschaft, stärken** möchte.[17] Diese wurde am 29. Juli 2020 durch vier Gründungsmitglieder zusammen mit der Europäischen Kommission ins Leben gerufen. Der Weg dahin begann aber bereits 2015, als die High Level Expert Group on the European Open Science

---

[11] Seite des NFDI, Seite zu Text+, https://www.nfdi.de/textplus/, abgerufen am 22.08.2023.
[12] Prof. Dr. Klaus Tochtermann im Podcast „The Future is Open Science” der ZBW mit Doreen Siegfried, https://podcast.zbw.eu/fos/folge-12-european-open-science-cloud/, zuletzt abgerufen am 22.08.2023.
[13] RatSWD-Seite des BAMF-FDZ, https://www.konsortswd.de/datenzentren/alle-datenzentren/bamf-fdz/, zuletzt abgerufen am 22.08.2023.
[14] RatSWD-Seite des FDZ RKI, https://www.konsortswd.de/datenzentren/alle-datenzentren/fdz-rki/, zuletzt abgerufen am 22.08.2023.
[15] Seite des NFDI, Seite zu FAIR Data Spaces, https://www.nfdi.de/fair-data-spaces/, zuletzt abgerufen am 22.08.2023.
[16] *Raum,* Ehmann/Selmayr, DSGVO. Art. 89, Rn. 6.
[17] *BMBF,* BMBF-Aktionsplan Forschungsdaten, 05. Oktober 2020.

Cloud erste Überlegungen hinsichtlich Anforderungen an ein europäisches Forschungsdatenmanagement angestellt hat. Zunächst förderte die Kommission einzelne Projekte, die unterschiedliche Aspekte im Bereich Forschungsdatenmanagement beleuchtete. Seit Gründung der European Open Science Cloud, einer Stiftung nach belgischem Recht, arbeitet diese daran, die Aktivitäten der Vergangenheit zu bündeln. Der Name EOSC hat jedoch mit dem, was sie ausmacht nichts zu tun. Denn die EOSC ist nicht europäisch, da „*Forschung nicht an regionalen Grenzen Europas endet, sondern viele Forschungsgruppen sind international vernetzt. Sie ist nicht Open, weil es auch in der Wissenschaft Daten gibt, die einen Schutzbedarf haben, beispielsweise Patientendaten. Sie ist nicht Science, weil viele wissenschaftliche Forschungsprojekte auch Daten aus der Wirtschaft nutzen. Und sie ist nicht Cloud, weil es nicht darum geht, alle Daten zentral in einer Cloudlösung abzulegen.*“[18] Gaia-X dagegen, die „*eine europäische Infrastruktur für das Management, also Speichern beispielsweise ihrer Daten*“ zum Gegenstand hat, hat die Wirtschaft als Zielgruppe. Die EOSC finanziert sich über Mitgliedsbeiträge, die laut Tochtermann 10.000 € für Vollmitglieder und 2.000 € für Beobachter beträgt. Im Jahr 2021 hatte die EOSC bereits 200 Mitglieder, woraus Tochtermann ein Jahresbudget von 1,5 Millionen Euro errechnete, um Personal für die Geschäftsstelle aufzubauen und die bislang Ehrenamtlichen des Board of Directors zu entlasten. Mitglieder sind keine Einzelpersonen, sondern Organisationen wie die ZBW oder der NFDI-Verein in Deutschland.[19]

2. Dat Act (COM(2022) 68 final)

Das Data Act schafft nicht nur allgemeine Datenzugangsansprüche der Benutzer und die Möglichkeit öffentlicher Stellen im Fall von „außergewöhnlichen Notwendigkeiten“ Daten von Dateninhabern zu verlangen. **Es besteht mit Art. 21 des Kommissionsentwurfs auch die Möglichkeit die dadurch „*erhaltenen Daten an Personen oder Organisationen zur Durchführung wissenschaftlicher Forschungstätigkeiten oder Analysen*“ weiterzugeben.** Diese Personen oder Organisationen müssen gemeinnützig sein oder gemäß einer anerkannten Aufgabe im öffentlichen Interesse handeln. Erwägungsgrund 68 weist in diesem Zusammenhang darauf hin, dass die Forschungstätigkeiten mit dem Zweck, wofür die Daten erlangt wurden, vereinbar sein und die Dateninhaber über die Weitergabe informiert werden müssen. Außerdem sollen Forschungsorganisationen, die einem erheblichen Einfluss gewerblicher Unternehmen unterliegen und die aufgrund der strukturellen Gegebenheiten Kontrolle ausüben können, nicht als Forschungseinrichtungen gelten. In Art. 21 Abs. 2 ist das Parlament von bestimmendem Einfluss auf erheblichen Einfluss ausgewichen.[20]

3. Verordnung über den europäischen Raum für Gesundheitsdaten

Für den Bereich der Gesundheitsdaten hat die Kommission am 03.05.2022 einen Vorschlag für einen gemeinsamen europäischen Datenraum veröffentlicht (COM(2022) 197 final). Teilweise in Anlehnung bzw. unter Berücksichtigung von Normen des DataAct wird natürlichen Personen das Recht eingeräumt auf ihre elektronischen Gesundheitsdaten zuzugreifen oder sie an Dritte weiterzugeben. Hierbei wird auch auf einen digitalen Zugang zu den personenbezogenen elektronischen Gesundheitsdaten und die Kontrolle darüber sowohl im eigenen Land als auch auf EU-Ebene wert gelegt. Der Entwurf unterscheidet zwischen Primär- (Kapitel II, Art. 3ff.) und Sekundärnutzung (Kapitel IV, Art. 33ff). Nach Art. 2 Abs. 2 lit. d ist **Primärnutzung** die Verarbeitung von personenbezogenen elektronischen Gesundheitsdaten für die Erbringung von Gesundheitsdiensten zur Beurteilung, Erhaltung oder Wiederherstellung des

[18] Prof. Dr. Klaus Tochtermann im Podcast „The Future is Open Science” der ZBW mit Doreen Siegfried, https://podcast.zbw.eu/fos/folge-12-european-open-science-cloud/, zuletzt abgerufen am 22.08.2023.
[19] Ebd.
[20] Änderungen des Parlaments zum Datengesetz, P9_TA(2023)0069, 14.03.2023.

Gesundheitszustands der natürlichen Person, einschließlich der Verschreibung, Abgabe und Bereitstellung von Arzneimitteln und Medizinprodukten sowie für die einschlägigen Dienste der Sozialversicherung, Verwaltung oder Kostenerstattung. Für die **Sekundärnutzung** werden zunächst Datenkategorien definiert (Art. 33) und sodann Verarbeitungszwecke festgelegt (Art. 34). Diese umfassen beispielsweise Tätigkeiten aus Gründen des öffentlichen Interesses im Bereich der öffentlichen Gesundheit wie Schutz vor schwerwiegenden grenzüberschreitenden Gesundheitsgefahren, die Unterstützung von öffentlichen Stellen im Gesundheits- oder Pflegesektor, Erstellung von Statistiken und Bildungs- oder Lehrtätigkeiten oder wissenschaftliche Forschung im Gesundheits- oder Pflegesektor (Art. 34 Abs. 1). **Art. 34 Abs. 1 lit. e des Entwurfs ist damit die Zweckeröffnung der Sekundärnutzung für die wissenschaftliche Forschung im Bereich des Gesundheits- oder Pflegesektors.** Art. 35 verbietet gewisse Zwecke der Sekundärnutzung, wie das Treffen von Entscheidungen mit Rechtswirkung zum Schaden einer natürlichen Person auf Grundlage ihrer elektronischen Gesundheitsdaten, Werbung oder die Entwicklung gesundheitsschädlicher Erzeugnisse oder Dienstleistungen. Zur Weitergabe der Daten sollen „Zugangsstellen für Gesundheitsdaten" (Art. 36ff.) eingerichtet werden. Hier haben Mitgliedstaaten die Möglichkeit mehrere Behörden einzurichten, in diesem Fall muss aber eine koordinierende Stelle eingerichtet werden. Die Zugangsstellen für Gesundheitsdaten prüfen nach Art. 46, ob Anträge von natürlichen oder juristischen Personen unter den Zweck des Art. 34 Abs. 1 fallen und Voraussetzungen des Kapitel IV erfüllen. Was unter wissenschaftlicher Forschung zu verstehen ist wird nicht definiert. Allerdings sagt Erwägungsgrund 41, dass die „*Sekundärnutzung von Gesundheitsdaten im Rahmen des EHDS […] es öffentlichen, privaten, nicht gewinnorientierten Einrichtungen sowie einzelnen Forschern ermöglichen [sollte], für die in der vorliegenden Verordnung festgelegten Zwecke Zugang zu Gesundheitsdaten für Forschung, Innovation, Politikgestaltung, Bildung, Patientensicherheit, Regulierungstätigkeiten oder personalisierte Medizin zu erhalten.*" Diese Nutzung soll dem allgemeinen Interesse der Gesellschaft dienen. Besondere Erwähnung findet auch die Unterstützung von Entwicklung und Innovation oder der Herstellung von Waren und Dienstleistung für den Gesundheits- oder Pflegesektor, wie das Trainieren von KI-Algorithmen, die die Gesundheit natürlicher Personen schützen oder Pflegetätigkeiten unterstützen könnten. Ein Ausschuss soll die Kommission und die Mitgliedstaaten unterstützen (Art. 64ff.).

4. Data Governance Act (Verordnung (EU) 2022/868)

Ein weiterer europäischer Rechtsakt, der die Nutzbarkeit von Daten verbessern möchte, ist der Data Governance Act. Hier geht es explizit um die Weiterverwendung von Daten in den Bereichen Gesundheit, Umwelt, Energie, Landwirtschaft, Mobilität, Finanzen, Fertigung, öffentliche Verwaltung und Qualifikation zum Vorteil von EU-Bürger:innen, Unternehmen, Innovation und der Schaffung von Arbeitsplätzen. Dabei stellt die Verordnung Regelungen für die Weiterverwendung bestimmter geschützter Daten, Vorschriften für Unternehmen, die Datenvermittlungsdienste erbringen, Vorschriften für Datenaltruismus, einen Europäischen Dateninnovationsraum und Maßnahmen für einen Verkehr nicht personenbezogener Daten außerhalb der EU auf.[21]

Einen **eigenen Datenraum** schafft der Rechtsakt nicht. Vielmehr wirkt er durch die Festlegung der einzelnen **Institutionen** wie den **Datenmittlern** und den **Vorgaben für Daten** öffentlicher Stellen auf einen **allgemeinen europäischen Datenraum** hin.

[21] Vgl. zusammenfassende Übersicht auf Eur-Lex.com zur Verordnung (EU) 2022/868.

5. Open Data und Public Sector Information Directive (Richtlinie (EU) 2019/1024)

Auch die OD-PSI kümmert sich um Forschungsdaten. Nach Art. 1 lit. c sind Forschungsdaten Gegenstand der Richtlinie. Art. 3 Abs. 1 stellt sodann generalisierend fest, dass grundsätzlich alle nach Art. 1 gegenständlichen Dokumente für kommerzielle und nichtkommerzielle Zwecke unter Beachtung der Kapitel III und IV weiterverwendet werden können. In Art. 1 und in Art. 3 Abs. 2 finden sich Einschränkungen hinsichtlich der Gegenstände. In Kapitel III finden sich Bedingungen für die Weiterverwendung. Insbesondere Formate, Gebühren, Lizenzen und in Art. 10 konkret zu Forschungsdaten. Art. 10 **greift die FAIR-Grundsätze auf und appelliert an die Mitgliedstaaten, nationale Strategien und Maßnahmen mit dem Ziel öffentlich finanzierte Forschungsdaten standardmäßig öffentlich zugänglich zu machen.** Kapitel IV enthält zwei Vorschriften zu Nichtdiskriminierung und lauterem Handel (Verbot von Ausschließlichkeitsvereinbarungen).

6. Free Flow of Data Act (Verordnung (EU) 2018/1807)

Der Free Flow of Data Act hat zwar nicht konkret Forschungsdaten zum Ziel, da er lediglich „*den freien Verkehr von Daten, die keine personenbezogenen Daten sind*“ (Art. 1) fördern möchte. Lediglich Erwägungsgrund 6 stellt fest, dass die Datenmobilität auch ein Problem für den Forschungs- und Entwicklungsbereich ist und Hochschulen und andere Forschungseinrichtungen mehr Zusammenarbeit brauchen, um die Innovationskraft zu stärken.

V. Datenzugangsansprüche

1. Datenzugang über Art. 31 Abs. 2 DSA-E bzw. Art. 40 Abs. 4 DSA

Im Digital Services Act (Verordnung (EU) 2022/2065, zuvor COM(2020) 825 final) findet sich „*ein originärer, subjektiv-rechtlicher Forschungsdatenzugangsanspruch gegen sehr große Online-Plattform- und Suchmaschinenanbieter*“[22], der „*sich aus dem Wortlaut von Art. 31 Abs. 2 DSA-E sowie […] Erwägungsgrund 64 DSA-E*“[23] ergibt.

„**Sehr große Online-**Plattformen gewähren auf **begründetes Verlangen** des Koordinators für digitale Dienste am Niederlassungsort oder der Kommission innerhalb einer darin genannten angemessenen Frist **zugelassenen Forschern**, die die Anforderungen in Absatz 4 dieses Artikels erfüllen, Zugang zu Daten zum **ausschließlichen Zweck** der Durchführung von **Forschungsarbeiten**, die zur **Ermittlung** und zum **Verständnis systemischer Risiken** gemäß Artikel 26 Absatz 1 beitragen.“[24]

Aufgrund der Tatsache, dass der Anspruch gegen die sehr große Online-Plattform bzw. -Suchmaschine nur über den Koordinator für digitale Dienste als Mittler durchsetzbar ist, sei es dogmatisch überzeugender „*von einem vermittelten Datenzugangsanspruch zu Gunsten der Forschung zu sprechen.*“[25]

Erwägungsgrund 64 ist mittlerweile Erwägungsgrund 96 des Textes, auf den sich Rat und Parlament am 21.09.2022 geeinigt haben (PE-CONS 30/22) und erläutert die dafür erforderlichen Zwecke bzw. Datenarten, die zu einer systemischen Analyse der Online-Risiken erforderlich sein können. Im angenommenen Text hat sich Art. 31 Abs. 2 DSA-E zu Art. 40 Abs. 4 verschoben. Dieser unterscheidet sich textlich dadurch, dass für die Aufspürung, Ermittlung und zum Verständnis der systemischen Risiken in der Union auch „*die **Bewertung** der **Angemessenheit**, der **Wirksamkeit** und der **Auswirkungen** der **Risikominderungsmaßnahmen** gemäß*

[22] *Wehde, Alexander,* Datenzugang über Art. 31 Abs. 2 DSA-E. In: MMR 2022, 827 (828).
[23] Ebd.
[24] Art. 31 Abs. 2 DSA-E der Kommission vom 15.12.2020.
[25] *Wehde, Alexander,* Datenzugang über Art. 31 Abs. 2 DSA-E. In: MMR 2022, 827 (828).

Artikel 35“[26] berücksichtigt werden soll und sehr große Online-Suchmaschinen nun auch textlich datenzugangsverpflichtet sind. Dieser Zugangsanspruch wird durch die Berechtigung als Forscher nach Art. 40 Abs. 8 der Verordnung bzw. Art. 31 Abs. 4 des Kommissionsentwurfs, der Zulässigkeit des Forschungsvorhabens und Zugangsgegenstands und möglichen entgegenstehenden Interessen und den faktischen Zugriffsmöglichkeiten der Anbieter beschränkt.

2. Datenzugang über § 19 Abs. 3 UrhDaG

Das die EU-Urheberrechterichtlinie umsetzende deutsche Urheberrechts-Diensteanbieter-Gesetz enthält in seinem § 19 Abs. 3 einen **Datenzugangsanspruch zum Zweck der wissenschaftlichen Forschung**. Er verweist auf Berechtigte nach § 60d Abs. 2 UrhG. Nach Ansicht von Prof. Specht-Riemenschneider ist § 19 Abs. 3 UrhDaG lediglich ein abgeleiteter Forschungsdatenzugangsanspruch, der sich an einen bestehenden Datenzugangsanspruch anlehnt. Sie begründet dies damit, dass bereits § 19 Abs. 2 einen bereits bestehenden Datenzugangsanspruch enthält.[27] Jan Oster dagegen hält Abs. 3 für eine eigene Anspruchsgrundlage der Berechtigten nach § 60d Abs. 2 UrhG und verweist begründend auf § 194 Abs. 1 BGB.[28] Bullinger sieht das Problem nicht und geht von einer Anspruchsqualität aus.[29] Letztlich soll § 19 Abs. 3 aber auch nach den Gesetzesmaterialien echten Anspruchscharakter haben, um **die Schaffung einer belastbaren empirischen Grundlage für die Debatte eines strukturellen Overblockings** zu ermöglichen.[30]

3. Datenzugang über § 5a NetzDG

Der Anspruch aus § 5a über den Einsatz und die **konkrete Wirkungsweise von Verfahren zur automatisierten Erkennung von Inhalten**, die entfernt oder gesperrt werden sollen oder die Verbreitung von Inhalten, die Gegenstand von Beschwerden über rechtswidrige Inhalten waren oder die vom Anbieter entfernt oder gesperrt wurden (§ 5a Abs. 2) steht grundsätzlich allen **natürlichen oder juristischen Personen** zu, die **wissenschaftliche Forschung** betreiben, § 5a Abs. 1. § 5a Abs. 3 verlangt einschränkend, dass die Auskünfte für eine im öffentlichen Interesse liegenden wissenschaftlichen Forschung zu Art, Umfang und Wirkungsweise öffentlicher Kommunikation in sozialen Netzwerken und den Umgang der Anbieter damit, erforderlich sind. Der Anbieter eines sozialen Netzwerks kann von dem Forscher **angemessene Kosten** für die Auskunftserteilung verlangen, höchstens jedoch 5.000 €. Bei einem außergewöhnlich hohen Aufwand darf dieser Betrag überschritten werden. Die Kostenerstattung darf kein Hindernis zur Geltendmachung des Auskunftsanspruchs sein.

4. Datenzugang über § 8 BKRG, § 13 Abs. 1 BayKRegG

Das Zentrum für Krebsregisterdaten soll öffentlichen und privaten Einrichtungen und Personen zu **wissenschaftlichen Forschungszwecken** Daten übermitteln. Hierzu müssen die Daten für das Vorhaben **geeignet und erforderlich** sein und überhaupt mit dem beim Zentrum vorliegenden Daten bearbeitet werden können und eine länderübergreifende Auswertung erfordern. Das Zentrum entscheidet unter Berücksichtigung einer Stellungnahme des wissenschaftlichen Ausschusses und übermittelt **grundsätzlich** nur **anonymisierte Daten**. Pseudonymisierte Daten sind nur mit besonderer Begründung der Erforderlichkeit und gesicherter physischer oder virtueller Umgebung unter Kontrolle des Zentrums für

[26] Art. 40 Abs. 4 DSA, angenommener Text PE-CONS 30/22, 22.09.2022.
[27] Dreier/Schulze, Urheberrechtsgesetz/Specht-Riemenschneider, 7. Aufl. 2022, UrhDaG, § 19 Rn. 15.
[28] BeckOK UrhR/Oster, 35. Ed. 15.7.2022, UrhDaG § 19 Rn. 8.
[29] Wandtke/Bullinger/Bullinger, 6. Aufl. 2022, UrhDaG § 19 Rn. 15ff.
[30] BT-Drs. 19/29894, 95.

Krebsregisterdaten möglich. Empfänger pseudonymisierter Daten müssen der Geheimhaltungspflicht des § 203 StGB unterliegen. Auch das Bayerische Krebsregistergesetz erlaubt die Nutzung anonymisierter Daten soweit ein berechtigtes, insbesondere wissenschaftliches Interesse glaubhaft gemacht wird, § 13 Abs. 1 S. 1 BayKRegG. In besonders begründeten Fällen können pseudonymisierbare oder personenidentifizierbare Daten übermittelt werden, § 13 Abs. 1 S. 2 BayKRegG. In diesem Fall muss allerdings die Zustimmung des Staatsministeriums eingeholt werden, § 13 Abs. 2 BayKRegG.

5. Datenzugang über § 303e, 363 SGB V

Zur Wahrung der Datentransparenz hat der Bund ein Forschungsdatenzentrum für die Gesetzliche Krankenversicherung eingerichtet. An dieses übermittelt insbesondere der Spitzenverband Bund der Krankenkassen nach § 303b Abs. 3 pseudonymisierte Daten. Der Spitzenverband wiederum erhält von allen Krankenkassen nach § 303b Abs. 1 SGB V u.a. Angaben zu Alter, Geschlecht, Wohnort, Angaben zum Versicherungsverhältnis, kosten- und Leistungsdaten, Vitalstatus und Sterbedatum. Das Forschungsdatenzentrum, dass u.a. die wissenschaftliche Erschließung der Daten fördern soll (§ 303d Abs. 1 Nr. 10), macht die Daten für unterschiedliche Zwecke (§ 303e Abs. 2) mehreren Nutzungsberechtigten (§ 303e Abs. 1) nach § 303e Abs. 3 anonymisiert und aggregiert zugänglich. Zu ersterem gehören insbesondere die Verbesserung der Versorgungsqualität, Planung von Leistungsressourcen, **Forschung (Analyse von Behandlungsabläufen oder des Versorgungsgeschehens)** oder Unterstützung politischer Entscheidungsprozesse. Zu letzterem gehören neben dem Spitzenverband Bund der Krankenkassen, insbesondere die Krankenkassen, Hochschulen und anerkannte Hochschulkliniken, öffentlich geförderte Forschungseinrichtungen oder der gemeinsame Bundesausschuss. Nach § 303e Abs. 4 können auch pseudonymisierte Einzeldatensätze bereitgestellt werden. Die Verfassungsmäßigkeit des Abs. 4 wurde zuweilen aus sozialdatenschutzrechtlichen Gesichtspunkten – auch vom BfDI und dem Bundesrat – angezweifelt, da hiermit bei seltenen Erkrankungen eine Re-Identifizierbarkeit möglich sei.[31] Das BVerfG hatte dies in einer Eilentscheidung aber nicht als deutlich genug den Nachteilen (Ziel des Gesetzgebers aufgeschoben und erschwert und Daten nicht für die Forschung verwendbar) überwiegend angesehen. Eine noch zu erhebende Verfassungsbeschwerde sei aber weder offensichtlich unzulässig noch offensichtlich unbegründet.[32] Auch die Daten der **elektronischen Patientenakte** können auf Antrag der Nutzungsberechtigten nach § 363 Abs. 4 i.Vm. 303e Abs. 2 Nr. 7, 8 SGB V diesen bereitgestellt werden.

6. Datenzugang nach § 1g Abs. 5, 63a Abs. 5 StVG

Nach § 1g Abs. 5 StVG ist das **Kraftfahrtbundesamt** berechtigt die **beim Halter erhobenen nicht personenbezogenen Daten für verkehrsbezogene Gemeinwohlzwecke** an Hochschulen und Universitäten, außeruniversitäre Forschungseinrichtungen und Behörden mit Forschungs-, Entwicklungs-, Verkehrsplanungs- oder Stadtplanungsaufgaben zugänglich zu machen. Der § 63a Abs. 5 StVG richtet sich gegen **Fahrzeuge mit hoch- oder vollautomatisierter Fahrfunktion** und erlaubt dem Halter die Übermittlung von ermittelten Positions- oder Zeitangaben bei der Aufforderung zur Übernahme der Fahrzeugsteuerung oder einer technischen Störung zum Zwecke der Unfallforschung an Dritte. Nach der aktuellen Rechtslage hat der Halter hier zwar die Pflicht die Daten zu übermitteln. Ihm fehlt es aber an der

[31] *Specht-Riemenschneider, Louisa,* Studie zur Regulierung eines privilegierten Zugangs zu Daten für Wissenschaft und Forschung durch die regulatorische Verankerung von Forschungsklauseln in den Sektoren Gesundheit, Online-Wirtschaft, Energie und Mobilität, im Auftrag des BMBF. August 2021, S. 59
[32] BVerfG, ZD 2020, 412 (413); Beschluss vom 19.3.2020 – 1 BvQ 1/20; dazu auch *Specht-Riemenschneider, Louisa,* Ebd.

technischen Kontrolle und der rechtlichen Befugnis auf die Daten zuzugreifen.[33] Hier könnten in Zukunft die Zugriffsrechte des Data Act Abhilfe schaffen.

VI. Empfohlenes Datenzugangsökosystem von Specht-Riemenschneider

Ausweislich ihrer Forschungsergebnisse sind Datenzugangsansprüche *„zwingend gemeinsam mit Datenzugangsinfrastrukturen zu entwickeln, um eine hinreichende Effektivität der Ansprüche zu gewährleisten. Ein wichtiger Baustein in diesen Datenzugangsinfrastrukturen sind Forschungsdatenzentren und Datentreuhänder, zu denen auch Personal Information Management-Systeme (PIMS) zählen. Datenzugangsinfrastrukturen und Datenzugangsansprüche bilden ein Datenzugangsökosystem.*“[34]

Zur besseren Visualisierung des Vorschlags eines Datenzugangsökosystem von Prof. Dr. Louisa Specht-Riemenschneider für den Gesundheitssektor wird auf die in ihrem Forschungsbericht vorhandenen Abbildungen hingewiesen.[35] Der Forschungsbericht ist öffentlich abrufbar.

Ihr Datenzugangsökosystem besteht aus einer zentralen Koordinierungsstelle und flexiblen Datentreuhandstrukturen. Die Datentreuhandstrukturen sind erforderlich, da sich die Einrichtung einer zentralen Koordinierungsstelle sowie einer Vielzahl an Forschungsdaten-Hubs nur auf spezifische Daten beziehen können. Die Datentreuhandstrukturen werden auch als „Data Clean Room“ bezeichnet.[36] Dort gelinge die Auswertung vertikal verteilter Daten besser.[37]

Ihrer Studie hat Specht-Riemenschneider auch einen **Gesetzesentwurf für ein Gesundheitsforschungsdatenzugangsgesetz** beigefügt. Dieser bringt wichtige Definitionen (§ 2), wie den „Broad Consent“, „Anonymisierung“, „Wissenschaft und Forschung“ oder „Datentreuhand“, schafft eine Koordinierungsstelle für den Forschungsdatenzugang (§ 3f), etabliert ein Data Governance (§ 5) und ein Research Ethics Board (§ 6) sowie Forschungsdaten-Hubs (§ 7). Die Koordinierungsstelle führt ein Verzeichnis privater Gesundheitsdienstleistungen für den Datenzugang und entscheidet über die Anträge (§ 8). § 9 regelt die Übermittlung von Daten in die Forschungsdaten-Hubs, der Datenzugang von Wissenschaft und Forschung findet sich in § 10. § 11 verpflichtet zur Erarbeitung von Standards durch einen Expertenrat für Anonymisierung und Pseudonymisierung und stellt die De-Anonymisierung unter Strafe.[38] Der Antrag auf Datenzugang nach § 10 kann nur bei Stellung im elektronischen Antragsportal (§ 12), dem Vorliegen eines Schutzkonzepts nach § 13 und der Nichtbeeinträchtigung von Rechten und Freiheiten anderer Personen nach § 15 gewährt werden. § 14 legt eine Entscheidungsfrist für die Koordinierungsstelle fest. § 16 ermöglicht die Erhebung von Gebühren nach der Datentransparenz-Gebührenverordnung. § 17 regelt die Nachnutzung und insbesondere den Ausschluss von der Antragsberechtigung für drei Jahre bei Verstößen. Der letzte Paragraf, § 18, enthält Regelungen zu Datentreuhandsystemen.[39]

---

[33] *Specht-Riemenschneider, Louisa*, Studie zur Regulierung eines privilegierten Zugangs zu Daten für Wissenschaft und Forschung durch die regulatorische Verankerung von Forschungsklauseln in den Sektoren Gesundheit, Online-Wirtschaft, Energie und Mobilität, im Auftrag des BMBF. August 2021, S. 60.
[34] Ebd. S. 14.
[35] Ebd. S. 149.
[36] Ebd. S. 148.
[37] Ebd. S. 149.
[38] Ein solcher wurde abseits von Specht-Riemenschneiders Vorschlag zur Gesundheitsdatenforschung bereits auch von Roßnagel und Geminn gefordert: *Roßnagel, Alexander/ Geminn, Christian,* Vertrauen in Anonymisierung. Regulierung der Anonymisierung zur Förderung Künstlicher Intelligenz. In: ZD 2021, 487 (490).
[39] *Specht-Riemenschneider, Louisa*, Studie zur Regulierung eines privilegierten Zugangs zu Daten für Wissenschaft und Forschung durch die regulatorische Verankerung von Forschungsklauseln in den Sek-

VII. Global Open Science Cloud

Die Global Open Science Cloud ist eine Initiative des „Masterminds“ hinter der Europäischen Open Science Cloud Jean-Claude Burgelman und dem Amerikaner George Strawn. Die GOSC besteht aus allen leitenden Figuren der regionalen Open Science Cloud-Bestrebungen. Das Ziel ist es frühzeitig die Interoperabilität der regionalen Datenmanagement-Lösungen, also u.a. der EOSC mit einer afrikanischen, australischen oder chinesischen OSC vorzubereiten.[40]

Auch CODATA (Comittee on Data of the International Science Council) unterstützt die GOSC und möchte die Zusammenarbeit der unterschiedlichen Open Science Plattformen stärken.[41]

toren Gesundheit, Online-Wirtschaft, Energie und Mobilität, im Auftrag des BMBF. August 2021, S. 150-157.

[40] Prof. Dr. Klaus Tochtermann im Podcast „The Future is Open Science” der ZBW mit Doreen Siegfried, https://podcast.zbw.eu/fos/folge-12-european-open-science-cloud/, zuletzt abgerufen am 22.08.2023.

[41] Seite von CODATA zur Global Open Science Cloud, https://codata.org/initiatives/decadal-programme2/global-open-science-cloud/ und Projektbericht zu DDI-CDI and EOSC, https://codata.org/initiatives/decadal-programme2/ddi-cross-domain-integration/ddi-cdi-and-eosc/, jeweils zuletzt abgerufen am 22.08.2023.