

Tätigkeitsbericht 2020

der vom RatSWD akkreditierten
Forschungsdatenzentren (FDZ)



Rat für Sozial- und Wirtschaftsdaten (RatSWD)

Tätigkeitsbericht 2020

der vom RatSWD akkreditierten
Forschungsdatenzentren (FDZ)

Inhaltsverzeichnis

Vorwort der Vorsitzenden des FDI Ausschusses	6
1 Überblick über die Forschungsdateninfrastruktur (FDI) des Rates für Sozial- und Wirtschaftsdaten (RatSWD)	7
Aktuelle Kennzahlen	7
Ausgangspunkt für die Etablierung des Netzwerks aus Forschungsdatenzentren	8
Weiterentwicklung und Sicherung der Forschungsdateninfrastruktur	8
Akkreditierungsprozess und Qualitätssicherung	9
Neue Akkreditierungen	10
Die gewachsene Infrastruktur in die Nationale Forschungsdateninfrastruktur integrieren	13
Herausforderungen in Zeiten der Covid-19 Pandemie	13
2 Struktur der Forschungsdatenzentren (FDZ)	14
Personal	14
Kooperationen und Forschungsaktivitäten	16
Wissenschaftliche Publikationen	17
3 Rotationsthema: Sicherung der Datenqualität	18
4 Datenangebot und Datennutzung	21
Umfang des Datenangebotes	21
Zeitpunkt der Verfügbarmachung der Datensätze und Gebühren	21
Forschungoutput auf Basis der bereitgestellten Forschungsdaten	23
Nutzungsumfang	24
5 Etablierte Datenzugänge und Servicekonzepte	26
Bewerbung des Datenangebotes	26
Zugangswege und -formate für die Datenbereitstellung	27
Bearbeitungsdauer nach Eingang eines unterzeichneten Vertrags	29
Bereitstellung von Arbeitshilfen	30
Serviceangebot für Datennutzende	31
Qualitätssicherung der Serviceangebote	33
6 Aktuelle Entwicklungen der Forschungsdateninfrastruktur in den Sozial-, Verhaltens- und Wirtschaftswissenschaften	34
Internationalisierung	34
Innovationen und Neuerungen zur bestehenden Dateninfrastruktur	35
Weiterentwicklung der Infrastruktur	36

7	Fokusthema Metadatenstandards	.37
	Findability (Auffindbarkeit)	.38
	Accessibility (Zugänglichkeit)	.40
	Interoperability (Interoperabilität)	.42
	Re-usability (Wiederverwendbarkeit)	.44
	FAZIT: FAIRe (Meta-)Daten?	.47
8	Beschwerdemanagement	.48
	Aktuelle Beschwerdeverfahren des Berichtsjahres 2020	.48

Anhang

	Anhang A: Entwicklung der Forschungsdateninfrastruktur (FDI) des RatSWD	.50
	Anhang B: Index und Angebote der Forschungsdateninfrastruktur (FDI) des RatSWD	.54
	Anhang C: Die Monitoringkommission	.60
	Anhang D: Mitwirkende bei der Erstellung des Tätigkeitsberichts 2020	.61

Vorwort der Vorsitzenden des FDI Ausschusses

Der Tätigkeitsbericht des Jahres 2020 verdeutlicht, wie erfolgreich die Forschungsdatenzentren (FDZ) mit den Herausforderungen der Pandemie umgehen konnten. Die Kennzahlen zu den erstellten Datenprodukten und den Bestellungen der Daten weisen darauf hin, dass das Kerngeschäft bei den meisten FDZ ohne Einschränkungen weitergelaufen ist. Hier zeigen sich die Vorteile einer weitgehenden Digitalisierung der Dienstleistungen einerseits und des hochmotivierten Personals andererseits, das sich auf neue Kommunikationswege und -formen schnell eingestellt hat. Auch die vielen Publikationen, die von den Beschäftigten erstellt wurden, sind ein Nachweis der Kreativität und Arbeitsfreude, die in den FDZ herrscht.

Das Feld der Daten, die durch die FDZ angeboten werden, wurde durch fünf weitere FDZ im Jahr 2020 erweitert. Es zeigt sich einerseits, dass neue Datenquellen für die Sozial- und Wirtschaftsforschung in Zukunft verbesserte wissenschaftliche Erklärungsmodelle ermöglichen werden. So kamen einerseits pharmazeutische Daten aus den Abrechnungen der Krankenkassen durch die Pharmakoepidemiologische Forschungsdatenbank (GePaRD) und andererseits Daten zu Siedlungsstruktur und Landnutzung durch den Monitor der Siedlungs- und Freiraumentwicklung (IÖR-Monitor) hinzu. Andererseits ergänzten Serviceeinrichtungen, welche sich der Verfügbarmachung vorhandener Daten aus Forschungsprojekten widmen, die Landschaft der FDZ. Sowohl das Forschungsdatenzentrum des Deutschen Zentrums für Integrations- und Migrationsforschung (DeZIM.fdz) als auch das Forschungsdatenzentrum für audio-visuelle Daten der qualitativen Sozialforschung (aviDa) erschließen Daten für die Nachnutzung und heben damit den Wert bereits vorhandener Datensätze. Eine Herausforderung bleibt die Integration europäischer Datensätze für eine über den nationalen Rahmen hinausreichende Forschung. Das Forschungsdatenzentrum des Leibniz-Instituts für Finanzmarktforschung SAFE ermöglicht in Zukunft auf dem Gebiet der Volkswirtschaft einen Zugang zu vereinheitlichten europäischen Daten. Auch nach dem Berichtsjahr 2020 setzte sich der Ausbau des Datenangebots durch zwei weitere FDZ im Jahr 2021 fort: Das Forschungsdatenzentrum der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (FDZ-BAuA) ermöglicht Zugang zu neuen Befragungsdaten zu den Themenfeldern Arbeit und Gesundheit; das Forschungsdatenzentrum des Bundesamtes für Migration und Flüchtlinge (BAMF-FDZ) stellt insbesondere Datenprodukte aus dem Ausländerzentralregister zur Verfügung.¹

Alle diese neuen FDZ sind herzlich willkommen in einem dynamischen Arbeitsfeld, das von ständigem Austausch lebt und in dem wir alle von Kooperation lernen und uns zu neuen Produkten und Serviceleistungen für die Forschung anregen lassen.

Dr. Daniel Fuß
Vorsitzender

Tatjana Mika
Vorsitzende

¹ Die FDZ GePaRD, aviDa, SAFE und BAMF-FDZ wurden vorläufig akkreditiert, da das operative Geschäft zunächst noch nicht aufgenommen wurde.

1 Überblick über die Forschungsdateninfrastruktur (FDI) des Rates für Sozial- und Wirtschaftsdaten

Der Ausschuss für Forschungsdateninfrastruktur (FDI Ausschuss) ist ein dynamisches dezentrales Netzwerk mit 39 vom Rat für Sozial- und Wirtschaftsdaten (RatSWD) akkreditierten Forschungsdatenzentren (FDZ), von denen fünf im Jahr 2020 neu akkreditiert wurden (Stand Dezember 2020).²

In den FDZ werden Daten archiviert und über verschiedene Zugangswege unter Einhaltung des Datenschutzes für die Wissenschaft zugänglich gemacht. Das Spektrum umfasst dabei etwa Studiendaten aus den verschiedenen Disziplinen der Sozial-, Verhaltens- und Wirtschaftswissenschaften ebenso wie Geo- und Raumdaten, Finanzdaten, statistische Daten von Bund und Ländern sowie Register- und Sozialversicherungsdaten. Nicht nur quantitative Daten, sondern auch qualitative Daten sind inzwischen zugänglich. Das Netzwerk wird beständig um weitere FDZ erweitert und damit auch um weitere Erhebungsmethoden sowie neue Datentypen und -formate angereichert. Der stetige Zuwachs an FDZ und steigende Nutzungszahlen verdeutlichen, dass das Modell zukunftsfähig ist und eine Akkreditierung durch den RatSWD gegenüber Drittmittelgebenden und Datennutzenden als Qualitätssiegel gilt und die empirische Forschung erleichtert.

Aktuelle Kennzahlen

Die Kennzahlen für das Jahr 2020 unterstreichen diese erfolgreiche Entwicklung der FDZ Landschaft: Zum Stichtag 31.12.2020 waren in den FDZ Mitarbeitende im Umfang von insgesamt 294 Vollzeitäquivalenten (VZÄ) beschäftigt, davon etwa zwei Drittel als wissenschaftliches Personal. Das FDZ Personal verfasste 2020 insgesamt 520 wissenschaftliche Publikationen. Eine positive Zahl in vielerlei Hinsicht: Die (Nach-)Nutzung der eigenen Forschungsdaten sichert den Anschluss an die Forschungscommunity, fördert beim FDZ Personal die Beratungskompetenz, dient als Qualitätskontrolle und steigert zugleich die Bekanntheit der Datensätze. Mindestens 2.906 wissenschaftliche Publikationen beruhen ganz oder teilweise auf den 4.917 von den FDZ bereitgestellten Datensätzen.

Für die Datennutzenden wurden 2020 insgesamt 546 neue Datensätze bereitgestellt – ein deutlicher Zuwachs im Vergleich zum Vorjahr.

Das Nachnutzungspotenzial von Forschungsdaten wächst. Von dem breiten Angebot der FDZ machten die Datennutzenden 2020 ausgiebig Gebrauch: Es gab insgesamt 43.703 Datennutzende in den FDZ und 68.752 freie Downloads wurden verzeichnet.



² Eine alphabetische Auflistung der FDZ sowie Angaben zum Datenangebot der verschiedenen FDZ (Kategorien: Sozial, Wirtschaft, Bildung, Gesundheit, Psychologie, Qualitativ, Weitere) findet sich in Anhang B.

Ausgangspunkt für die Etablierung des Netzwerks aus Forschungsdatenzentren³

Auf Initiative der Wissenschaft wurde 1999 die „Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik“ (KVI) gegründet.⁴ Dieses Gremium legte im März 2001 ein umfassendes Gutachten vor. Eine der zentralen Empfehlungen war die Einrichtung von FDZ bei den Statistischen Ämtern des Bundes und der Länder, der Deutschen Rentenversicherung Bund und der Bundesagentur für Arbeit, damit die dort vorhandenen Forschungsdaten professionell archiviert und soweit möglich etwa für die Replikation oder für die Beantwortung neuer Fragestellungen genutzt werden können. Im selben Jahr wurde der Gründungsausschuss des RatSWD konstituiert. Damit war der Grundstein für das dynamische Netzwerk rund um den RatSWD gelegt.

Weiterentwicklung und Sicherung der Forschungsdateninfrastruktur

Seit 2001 entwickelte sich die Forschungsdateninfrastruktur stetig weiter. In diesem Zusammenhang gründeten sich auch neue FDZ, wobei die Motive zur Gründung eines FDZ dabei sehr heterogen waren: Die einen folgten externen Empfehlungen, wie solchen der Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik, des Wissenschaftsrates oder wissenschaftlicher Beratungskreise. Andere erhielten aus der eigenen Einrichtung den Auftrag zur Gründung eines FDZ zum Zweck der Forschungsförderung. Allen gemein war jedoch das Bestreben, die Forschungsdateninfrastruktur in Deutschland durch die Bereitstellung von Forschungsdaten für die Wissenschaft zu erweitern und zu stärken.

Im Jahr 2004 wurde der RatSWD als strategisches Gremium gegründet. Um den operativen Dialog zwischen den FDZ zu fördern, wurde im Jahr 2009 der „Ständige Ausschuss Forschungsdateninfrastruktur“ (FDI Ausschuss) eingerichtet. Hauptaufgaben dieses Gremiums sind die kontinuierliche Sicherung und nachhaltige Verbesserung der Forschungsdateninfrastruktur: Dies beinhaltet den Ausbau der Qualität und Quantität des Datenangebots ebenso wie die Weiterentwicklung und Vereinfachung des Datenzugangs für die Forschungscommunity.

Infokasten 1

Aufgaben und Struktur des RatSWD

Der Rat für Sozial- und Wirtschaftsdaten (RatSWD) berät seit 2004 die Bundesregierung und die Regierungen der Länder in Fragen der Erweiterung und Verbesserung der Forschungsdateninfrastruktur für die empirischen Sozial-, Verhaltens- und Wirtschaftswissenschaften. Er ist ein institutionalisiertes Forum des Austausches und des Dialoges zwischen Wissenschaft und Datenproduktion über die Verbesserung des Zugangs zu qualitativ hochwertigen und wissenschaftlich interessanten Daten. Diese stammen aus den Tätigkeitsbereichen von staatlichen, wissenschaftsgetragenen und privatwirtschaftlichen Akteuren. Im RatSWD arbeiten (Stand der 7. Berufungsperiode ab 2020) zehn durch Wahl legitimierte Vertretende der sozial-, verhaltens- und wirtschaftswissenschaftlichen Fachdisziplinen sowie zehn Vertretende der wichtigsten Datenproduzierenden zusammen. Mit seiner paritätischen Besetzung ist der RatSWD ein fachkundiges und gleichzeitig breit aufgestelltes Gremium. Es spielt eine tragende Rolle bei der Weiterentwicklung der Forschungsinfrastrukturen für die empirischen Sozial-, Verhaltens- und Wirtschaftswissenschaften und setzt sich für eine wissenschaftsfreundliche Gestaltung der rechtlichen und politischen Rahmenbedingungen ein.

³ In Anhang A werden die Entwicklung und die Angebote der Forschungsdateninfrastruktur (FDI) des RatSWD chronologisch übersichtlich dargestellt.

⁴ Zapf, W. (1996). Memorandum zur Verbesserung der Zugangsmöglichkeiten zu Mikrodaten der amtlichen Statistik. *ZUMA-Nachrichten*, 39, 172-175; Hauser, R., Wagner, G. G., & Zimmermann, K. F. (1998). Erfolgsbedingungen empirischer Wirtschaftsforschung und empirisch gestützter wirtschafts- und sozialpolitischer Beratung. *Allgemeines Statistisches Archiv*, 82(3), 369-379.

Akkreditierungsprozess und Qualitätssicherung

Zur Qualitätssicherung der Forschungsdateninfrastruktur hat der RatSWD im Jahr 2010 Mindeststandards und Kriterien für die Akkreditierung erarbeitet.⁵ Seit 2015 werden diese laufend an technische Neuerungen und methodische Erweiterungen des Netzwerkportfolios angepasst. Durch ein jährliches Monitoring werden die Standards überprüft und die Qualität des Angebots gesichert. Die Sichtung der Akkreditierungsunterlagen neuer FDZ sowie das jährliche Monitoring werden von einer Monitoringkommission realisiert, die aus dem FDI Ausschuss gewählt wird.⁶ Über eine Akkreditierung entscheidet abschließend der RatSWD. Zusätzlich nehmen die FDZ an einem Beschwerdemanagement teil, das im Falle von nicht bilateral zu lösenden Fragen in der Datenbereitstellung bei der Geschäftsstelle des RatSWD administriert und von der Monitoringkommission verantwortet wird (vgl. Kapitel 8).

Abb. 1: Zusammenarbeit von FDI Ausschuss und RatSWD



Zusammengefasst werden die Ergebnisse des Monitorings in einem internen und dem hier vorliegenden öffentlich zugänglichen jährlichen Tätigkeitsbericht der vom RatSWD akkreditierten FDZ.⁷ Die Akkreditierung durch den RatSWD ist für alle FDZ ein Qualitätssiegel, denn sie setzt die Einhaltung von Pflichtkriterien voraus: So muss zumindest ein Datenzugangsweg existieren, Datensätze müssen hinreichend dokumentiert sein und es muss ein Konzept zur längerfristigen Verfügbarmachung der Daten vorgelegt werden. Ausschlaggebend für die Akkreditierung sind außerdem Informationen, etwa über vorhandene Arbeitshilfen, Qualitätssicherung und Angaben zur Weiterentwicklung der Infrastruktur und zur Sicherstellung des Datenschutzes.⁸ Die Akkreditierung zahlt sich für die FDZ in verschiedener Hinsicht aus: Die FDZ werden infolge der Akkreditierung in ihrer Arbeit unterstützt, werden für die Weiterentwicklung der eigenen Infrastrukturen über Best Practice-Lösungen informiert und partizipieren am Erfahrungs- und Wissensaustausch mit den weiteren FDZ.

5 RatSWD [Rat für Sozial- und Wirtschaftsdaten] (2017). *Qualitätssicherung der vom Rat für Sozial- und Wirtschaftsdaten (RatSWD) akkreditierten Forschungsdatenzentren (FDZ)*. RatSWD Output 8(5). <https://doi.org/10.17620/02671.4>.

6 Hintergrundinformationen und die Mitglieder der Monitoringkommission sind in Anhang C zu finden.

7 Die Tätigkeitsberichte ab dem Jahr 2015 sind abrufbar unter: <https://www.konsortswd.de/aktuelles/publikationen/taetigkeitsberichte> (Zugriff am 31.01.2022).

8 Vgl. zu den Akkreditierungskriterien RatSWD [Rat für Sozial- und Wirtschaftsdaten] (2017). *Qualitätssicherung der vom Rat für Sozial- und Wirtschaftsdaten (RatSWD) akkreditierten Forschungsdatenzentren (FDZ)*. S. 6 f. RatSWD Output 8(5). <https://doi.org/10.17620/02671.4>.



Neue Akkreditierungen

Im Jahr 2020 wurden fünf weitere FDZ akkreditiert. Die neuen FDZ erweitern den Kreis der akkreditierten FDZ maßgeblich. So werden medizinisch-epidemiologische (GePaRD), finanzwirtschaftliche (SAFE FDZ) und raumbezogene (IÖR-Monitor) Daten, sowie Daten zu Migration und Integration (DeZIM.fdz) für die sozial-, verhaltens- und wirtschaftswissenschaftliche Forschung zugänglich. Mit dem FDZ-aviDa gibt es nun erstmals ein Angebot für video-grafische Forschungsdaten.⁹

Pharmakoepidemiologische Forschungsdatenbank (GePaRD)

Das FDZ GePaRD (German Pharmacoepidemiological Research Database) ist eine pharmakoepidemiologische Forschungsdatenbank basierend auf Daten gesetzlicher Krankenversicherungen in Deutschland, welche seit 2004 am Leibniz-Institut für Präventionsforschung und Epidemiologie – BIPS aufgebaut und kontinuierlich gepflegt wird. Mit GePaRD können u. a. Forschungsfragen zum Gebrauch und zur Sicherheit von Arzneimitteln und Impfstoffen in der Routineversorgung untersucht werden.

GePaRD enthält Abrechnungsdaten von vier gesetzlichen Krankenversicherungen und umfasst Informationen von derzeit ungefähr 25 Millionen Personen, die seit 2004 oder später bei einer der teilnehmenden Krankenversicherungen versichert waren. Neben demographischen Angaben enthält GePaRD Informationen zu erstattungsfähigen Arzneimittelverordnungen sowie zu ambulanten und stationären Leistungen und Diagnosen. Pro Datenjahr stehen Informationen zu ungefähr 20 % der Allgemeinbevölkerung zur Verfügung und es sind alle geographischen Regionen Deutschlands vertreten. Der gesamte Prozess von der Datenlieferung über die Datenaufbereitung bis zur Verfügbarkeit der Daten für Auswertungen dauert meist 24 Monate, d. h. beispielsweise, dass Daten aus dem Jahr 2018 frühestens Ende 2020 verwendet werden können.

Das BIPS ist nicht Eigentümer der Daten und darf daher nicht entscheiden, für welche konkreten Projekte die Daten verwendet werden können. Die Genehmigung von Projekten basiert auf der Bewilligung des Vorhabens durch die Krankenversicherungen und die jeweils zuständigen Behörden. Entscheidend für die Nutzung der Daten gemäß § 75 SGB X ist, dass das öffentliche Interesse das schutzwürdige Interesse der betroffenen Personen erheblich überwiegt. Der Prozess der Genehmigung durch die Krankenversicherungen und die zuständigen Behörden dauert in der Regel mindestens drei Monate.

Forschungszentrum des Leibniz-Instituts für Finanzmarktforschung SAFE

Die wirtschaftswissenschaftliche Forschung im deutschen und europäischen Kontext wird dadurch erschwert, dass einerseits wenig paneuropäische Datensätze existieren und andererseits in den Datenbeständen bisher keine einheitliche Identifikation von z. B. Firmen möglich ist. Deshalb weichen Forschende oftmals auf US-amerikanische Daten aus, in denen die Integration verschiedener Datenbanken schon weiter fortgeschritten ist. Das hat aber zur Folge, dass empirische Analysen oftmals nur mithilfe nicht-europäischer Datensätze durchgeführt werden, sodass die daraus abgeleiteten Politikempfehlungen aufgrund der institutionellen Unterschiede nicht oder nur eingeschränkt auf den europäischen Raum übertragbar sind.

Vor diesem Hintergrund greift das SAFE-FDZ nicht nur auf die üblichen internationalen Datenquellen zurück, sondern erhebt auch neue europäische Datensätze, führt vorhandene zusammen und bereitet sie auf. Ziel ist es, die fünf zentralen Forschungsbereiche von SAFE („Financial Intermediation“, „Financial Markets“, „Household Finance“, „Macro Finance“ und „Money & Finance“) auf eine gemeinsame europäische Datenbasis zu stellen.

⁹ Die FDZ GePaRD, aviDa und SAFE wurden vorläufig akkreditiert, da das operative Geschäft zunächst noch nicht aufgenommen wurde.



Monitor der Siedlungs- und Freiraumentwicklung (IÖR-Monitor)

Der IÖR-Monitor ist eine Forschungsdateninfrastruktur des Leibniz-Instituts für ökologische Raumentwicklung (IÖR). Er stellt Daten und Informationen zur Flächennutzungsstruktur und deren Entwicklung sowie zur Landschaftsqualität für die Bundesrepublik Deutschland bereit.

Die flächenbezogenen Daten stehen hochaufgelöst auf geographischen Gitterzellen sowie unterschiedlichen administrativen Ebenen mit jährlicher Aktualität bereit. Die Datensätze basieren auf Forschungsergebnissen des IÖR. Kennblätter informieren als Metadaten über Datengrundlagen und Berechnungsmethodik. Grundlage für die Berechnung sind insbesondere geotopographische Daten (ATKIS Basis-DLM), Daten zur Landbedeckung (LBM-DE), amtliche Hausumringe (HU-DE) und Hauskoordinaten (HK-DE) sowie weitere Geofach- und Statistikdaten.

Die Datennutzung wird über einen Kartenviewer mit umfangreichen Interaktionsmöglichkeiten, einen Datenexport und Geodienste angeboten. Die Daten können von interessierten Wissenschaftler:innen für bestimmte räumliche Ausdehnungen und Zeiträume bedarfsgerecht bezogen werden. Ein Überblick über die mehr als 80 Datensätze samt Berechnungsmethoden, räumlichem und zeitlichem Bezug und den entsprechenden Exportfunktionen mittels Geodiensten ist unter <https://www.ioer-monitor.de/indikatoren> möglich.



Forschungsdatenzentrum des Deutschen Zentrums für Integrations- und Migrationsforschung (DeZIM.fdz)

Das Deutsche Zentrum für Integrations- und Migrationsforschung (DeZIM) ist eine bundesweite Initiative von Wissenschaft und Politik zur Stärkung einer exzellenten und international sichtbaren Integrations- und Migrationsforschung in Deutschland. Das DeZIM.fdz bietet interessierten Wissenschaftler:innen die Möglichkeit, auf Forschungsdaten zuzugreifen, die im Zuge der am DeZIM-Institut sowie an den Instituten der DeZIM-Gemeinschaft durchgeführten Projekte erhoben wurden. Neben dem Zugang zu den Daten bietet das DeZIM.fdz eine umfangreiche Informationsdatenbank an, die eine Recherche zu – intern wie auch extern archivierten – Studien ermöglicht, die sich mit den Themenbereichen Migration und Integration befassen. Weiterhin organisiert das DeZIM.fdz einen umfangreichen Support zu den einzelnen Datensätzen in seinem Datenangebot sowie zu verschiedenen methodischen Schwerpunktthemen.



Forschungsdatenzentrum für audio-visuelle Daten der qualitativen Sozialforschung (aviDa)

Das Angebot von aviDa zielt darauf ab, von Forschenden selbst erhobene videographische Forschungsdaten zu teilen und für die Nachnutzung zu öffnen. Es richtet sich an alle videografisch Forschenden im Feld der qualitativen empirischen Sozialforschung. Auf aviDa können bei Primärforschung erhobene Forschungsdaten Dritten zugänglich gemacht und für den Einsatz in Forschung und Lehre webbasiert nachgenutzt werden.

Auf Basis von DepositOnce, dem Repository für Forschungsdaten und Publikationen der Technischen Universität Berlin, wurde aviDa in Kooperation des Fachgebiets Allgemeine Soziologie der Technischen Universität Berlin mit dem Lehrstuhl für Kultur- und Religionssoziologie der Universität Bayreuth, der Universitätsbibliothek sowie der Zentraleinrichtung Campusmanagement (ZECM) der TU Berlin seit Juli 2018 als prototypische Forschungsdateninfrastruktur (FDI) für die langfristige, Community-basierte digitale Sicherung und Nachnutzung audio-visueller Forschungsdaten aufgebaut. In einem ersten Projektteil wurde der Grundstein für eine nachhaltige und stabile, webbasierte Plattform mit hoher Usability und Effizienz gelegt, die dem (wissenschaftsinternen) Austausch von Forschungsdaten zur Nachnutzung in Forschung und Lehre dienen soll.

Abb. 2:
**Standorte und Gastwissenschaftsarbetsplätze der
akkreditierten Forschungsdatenzentren in 2020**

Das FDZ BA im IAB verfügt zudem über Standorte mit Gastwissenschaftsarbetsplätzen in Europa, Kanada und den USA.



- Forschungsdatenzentren (FDZ)
- FDZ mit Gastwissenschaftsarbetsplätzen
- externe Gastwissenschaftsarbetsplätze eines FDZ in einem anderen FDZ oder einer anderen Einrichtung

Die gewachsene Infrastruktur in die Nationale Forschungsdateninfrastruktur integrieren

Im Oktober 2020 hat die Arbeit des Konsortiums für die Sozial-, Verhaltens-, Bildungs- und Wirtschaftswissenschaften (KonsortSWD) im Rahmen der Förderung der Nationalen Forschungsdateninfrastruktur (NFDI) begonnen. Die Dienste, die durch das Konsortium entwickelt werden, sollen in Zukunft die Leistungen der FDZ weiter verbessern. Deshalb werden die FDZ in alle Arbeitspakete eingebunden und das Konsortium berichtet von nun an regelmäßig dem FDI Ausschuss über den Projektfortschritt. Für die FDZ besonders relevant sind die Aktivitäten zur Professionalisierung des Forschungsdatenmanagements und der verstärkten Vernetzung der FDZ Angebote. Direkte Unterstützung erhalten die FDZ für die Gestaltung von Richtlinien und Verträgen, aber auch für die (Weiter-)Qualifikation der Mitarbeitenden. Aus Sicht der Nutzenden ist die Vernetzung von Gastwissenschaftsarbeitsplätzen (GWAP) eine wichtige Maßnahme, die dabei helfen wird, Aufwände für den Datenzugang zu verringern. Eine Übersicht über die Dienste des Konsortiums findet sich auf der Webseite: <https://www.konsortswd.de/services>.



Herausforderungen in Zeiten der COVID-19-Pandemie

Die COVID-19-Pandemie stellte die FDZ im Jahr 2020 (und auch darüber hinaus) vor besondere Herausforderungen, vornehmlich in den Bereichen Datenzugang, Vertragsgestaltung sowie in der Organisation und Durchführung von Workshops oder Schulungen. Beim Datenzugang waren aufgrund der Kontaktbeschränkungen insbesondere die GWAP betroffen, die während des Pandemieerlaufs geschlossen werden mussten. Die Öffnung von Gasträumen hing nicht nur vom regionalen Pandemiegeschehen in 2020 selbst und der Festlegung geeigneter Maßnahmen zur Ansteckungsreduktion ab, sondern auch von Entscheidungen der jeweiligen Einrichtungen zum Schutz der Belegschaft. Einzelne FDZ schafften Überbrückungsmöglichkeiten wie z. B. die Datenfernverarbeitung als alternativen Datenzugangsweg. Auch die Nutzung von Scientific Use Files ist vertraglich oft nur in der jeweiligen wissenschaftlichen Einrichtung der Nutzenden und nicht von zuhause aus per sicherem Remote-Verfahren möglich. Hier ermöglichten FDZ zeitlich befristete Ausnahmeregelungen wie beispielsweise Neuregelungen beim Datenzugang im Homeoffice. Bei der Vertragsgestaltung führte insbesondere die notwendige Einholung von Originalunterschriften zu zeitlichen Verzögerungen beim Datenzugang und konnte beispielsweise durch die Einführung von zertifizierten Signaturen abgefedert werden. Wie schnell FDZ situativ Änderungen vornehmen, zeigt sich auch beim Anstieg von virtuellen Schulungen und Workshops, die vorher in Präsenz stattfanden.

2 Struktur der Forschungsdatenzentren (FDZ)

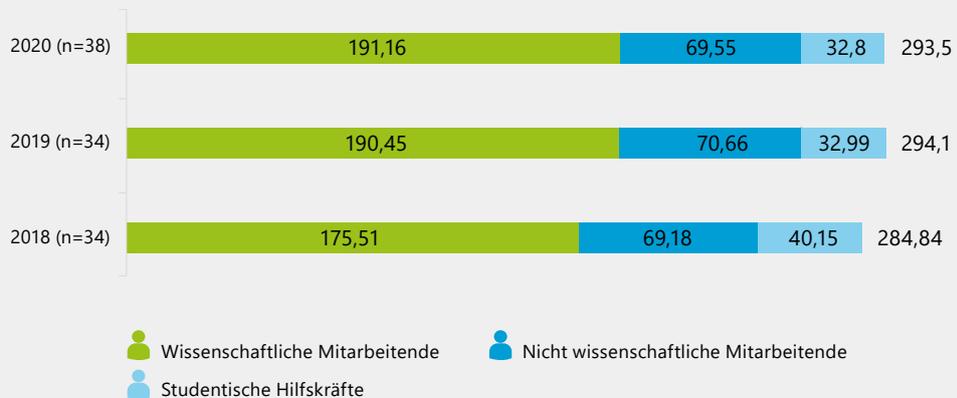
Die in den folgenden Kapiteln zusammengefassten Informationen entstammen dem jährlichen Monitoring, das im Austausch mit allen FDZ (weiter-)entwickelt wurde und an dem sich entsprechend auch alle vom RatSWD akkreditierten FDZ beteiligen. Im Berichtsjahr 2020 nahmen 39 FDZ am Monitoring teil. Dabei erfolgte die Beantwortung der Fragen durch die FDZ der Statistischen Ämter des Bundes und der Länder gemeinsam, sodass die folgenden Angaben sich auf 38 Antworten aus den FDZ beziehen.

Personal

Die FDZ hatten ihre personelle Ausstattung in den vergangenen Jahren kontinuierlich erweitert. Nunmehr ist eine Konsolidierungsphase zu verzeichnen. Der Personalbestand der FDZ stabilisiert sich auf einem hohen Niveau. Zwar sind im Jahr 2020 vier neue FDZ dazu gekommen, was aber nicht zu einem weiteren Personalaufwuchs geführt hat. Zum 31.12.2020 beschäftigten die 38 FDZ Personal in der Stärke von 293,5 Vollzeitäquivalenten (VZÄ). Der Aufwuchs beim Personal zeigte sich in den Vorjahren vor allem im wissenschaftlichen Bereich. Die Anteile des nicht wissenschaftlichen Personals und der studentischen Hilfskräfte befinden sich ebenfalls nahezu unverändert zum Vorjahr (► Abb. 3). Ein gutes Drittel der FDZ arbeitet gänzlich ohne studentische Unterstützung.

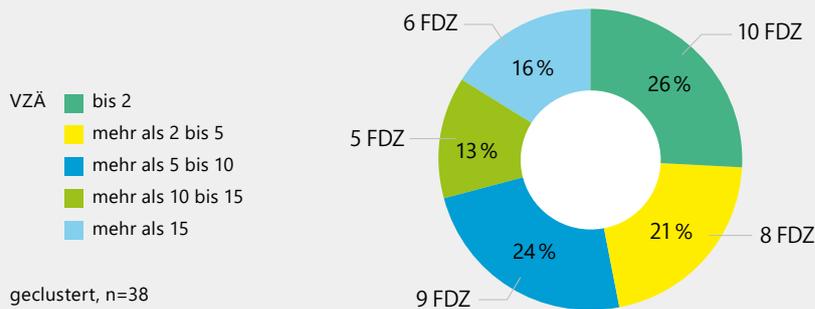
Abb. 3: Personal und Verteilung des Personalbestands in Vollzeitäquivalenten (VZÄ)

Der Personalzuwachs der Vorjahre hat sich nicht fortgesetzt.



Die Anzahl der Mitarbeitenden in den FDZ variiert stark. Im Mittel verfügt ein FDZ über 7,7 Vollzeitäquivalente, tatsächlich reicht die Spannweite von weniger als einem VZÄ bis zu 34 VZÄ. Große FDZ sind aber eher selten. Es dominieren die kleineren Einheiten. Knapp drei Viertel der FDZ haben in VZÄ maximal zehn Mitarbeitende (► **Abb. 4**). Falls das FDZ zu einer Institution gehört, so lässt sich kein unmittelbarer Zusammenhang zur Personalstärke dieser Institution herstellen.

Abb. 4: Personalbestand in Vollzeitäquivalenten (VZÄ)

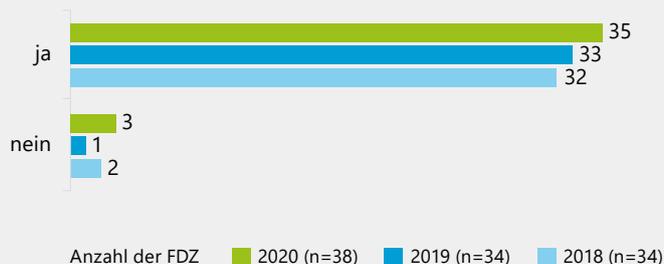


Die Personalausstattung in den FDZ variiert stark, liegt aber bei den meisten FDZ bei fünf oder mehr VZÄ

In 35 der 38 befragten FDZ betreibt das wissenschaftlich ausgebildete Personal Forschung (► **Abb. 5**). Diese Forschungsaktivitäten können inhaltliche Beiträge zur Forschungsagenda zum Thema haben, sich aber auch methodisch mit Daten befassen oder sich auf technologisch-funktionelle Weiterentwicklungen der Dateninfrastruktur fokussieren. Von diesen Aktivitäten profitieren letztlich auch die Datennutzenden in Form von kompetenter Beratung auf aktuellem Wissensstand.

Abb. 5: Selbstständige Forschung von wissenschaftlichem Personal

Gibt es wissenschaftlich ausgebildetes Personal am FDZ, das selbstständig forscht?



In nahezu allen FDZ forschen wissenschaftlich ausgebildete Mitarbeitende auch selbst

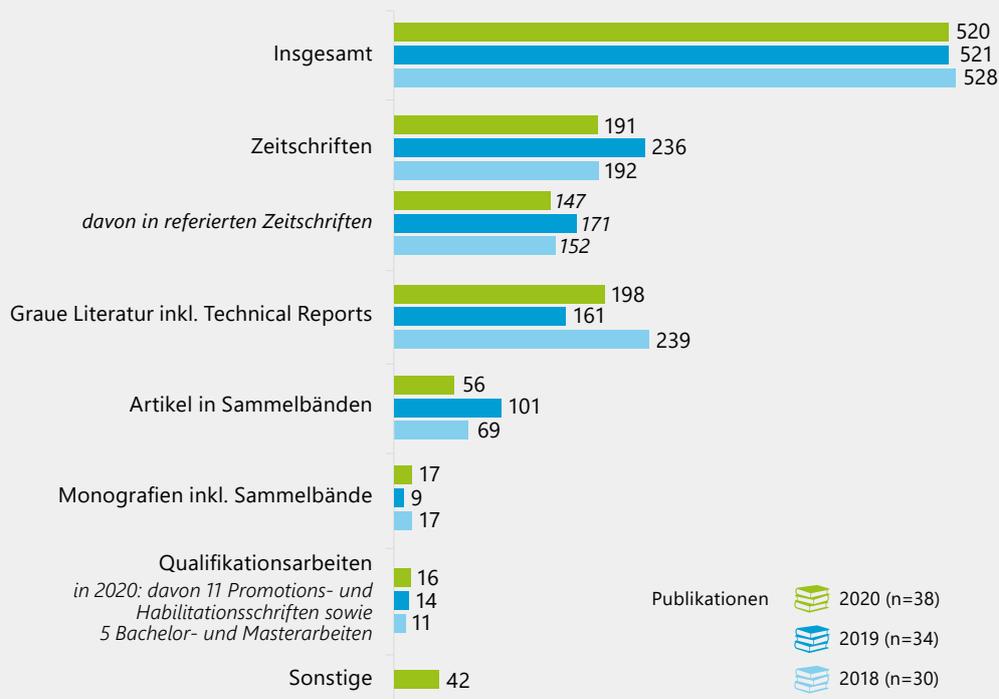
Zum Umfang dieser Tätigkeit: Feste Arbeitszeitanteile für eigene Forschung des wissenschaftlichen Personals gibt es in 17 FDZ. Die Anteile betragen im Durchschnitt ein knappes Drittel der Arbeitszeit, mit einer breiten Streuung in den einzelnen FDZ. In sieben FDZ wird bis zu einem Viertel der Arbeitszeit dafür aufgewendet, in zehn FDZ liegt der Anteil darüber.

Wissenschaftliche Publikationen

Wie bereits dargestellt, besteht die Mehrheit der Vollzeitäquivalente in den FDZ aus wissenschaftlichen Mitarbeitenden und diese forschen in der Regel auch selbst. Die Ausrichtung der FDZ ist nicht nur auf Serviceleistungen und Bereitstellung von Daten fokussiert. Sie schließt auch qualifizierte Beratung zu den Potenzialen bereitgestellter Datensätze ein. Diese muss nicht auf inhaltliche Fragestellungen begrenzt sein, sondern umfasst beispielsweise auch Beratung zu Anwendungsmöglichkeiten und Restriktionen statistischer Verfahren. Eigene Forschung zu und mit den angebotenen Daten trägt dazu bei, dass die Daten-, Service- und Beratungstätigkeiten stets auf angemessenem wissenschaftlichem Niveau durchgeführt werden. Des Weiteren wird sie auch in wissenschaftlichen Publikationen dokumentiert. Insgesamt ist der Publikationsoutput 2020 in etwa auf Vorjahresniveau geblieben, wie ► **Abb. 7** zeigt. Im Berichtszeitraum wurden 520 Publikationen vom FDZ Personal verfasst.

Abb. 7: Wissenschaftliche Publikationen des FDZ Personals

Bitte geben Sie die Anzahl der wissenschaftlichen Publikationen Ihrer FDZ Mitarbeitenden an, unabhängig von den verwendeten Daten und unabhängig davon, ob die Publikation im Rahmen der FDZ Arbeitszeit verfasst wurde.



Das FDZ Personal veröffentlichte viele Beiträge in referierten Zeitschriften

478 Veröffentlichungen lassen sich in die wichtigsten wissenschaftlichen Publikationsformate einordnen. Wie schon in den Vorjahren sind die Beiträge am häufigsten in der Grauen Literatur und in Zeitschriften veröffentlicht worden. Es ist auch wieder gelungen, sehr viele Artikel in referierten Zeitschriften zu veröffentlichen, was als Ausweis der hohen Qualität der Inhalte gewertet werden kann. Die Zahl der Qualifikationsarbeiten ist leicht gestiegen. Es handelt sich analog zu den Vorjahresergebnissen vor allem um höhere Qualifikationen (Promotion, Habilitation). Bachelor- und Masterarbeiten zählen ebenfalls zu den Qualifikationsarbeiten, werden aber nicht von jedem FDZ erfasst, weshalb an dieser Stelle von einer Untererfassung ausgegangen werden kann. Immerhin 42 Veröffentlichungen fallen in die Kategorie „Sonstige“, in die u. a. Tagungsbände und Variablenreports einzuordnen sind.

3 Rotationsthema: Sicherung der Datenqualität

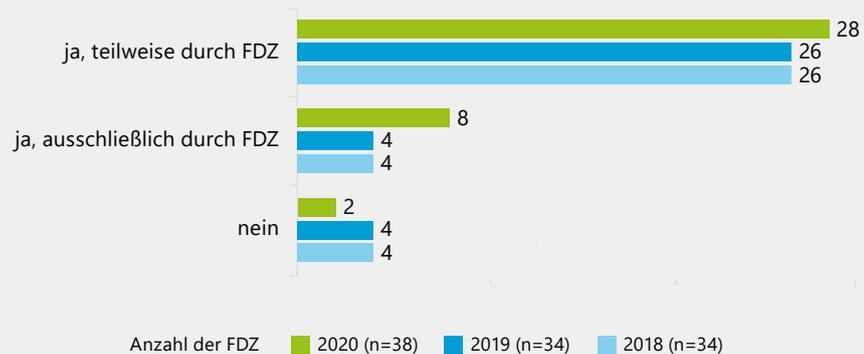
Zu den Aufgaben der FDZ gehören vor allem die langfristige Sicherstellung von Forschungsdaten und deren Dokumentation zur Nachnutzung durch Dritte, die Sicherstellung des Datenschutzes sowie in den meisten FDZ zusätzlich die vorherige Prüfung der Daten auf ihre Qualität und Güte. Während in der Vergangenheit alle drei Bereiche für den jährlichen Tätigkeitsbericht erhoben wurden, werden die Bereiche ab dem vorliegenden Tätigkeitsbericht 2020 auf drei Erhebungsjahre verteilt.¹⁰ Für das Jahr 2020 wurde das Vorgehen der FDZ zur Qualitätssicherung der Datensätze abgefragt.

Die Einbindung der FDZ in die Qualitätssicherung der Daten hängt u. a. damit zusammen, ob und wie ein FDZ mit einer weiteren (Träger-)Institution verbunden ist. Die Art und Weise des Einbezugs der FDZ in die Sicherung der Datenqualität ist üblicherweise über die Zeit stabil. So gab es kaum eine Änderung in der Aufgabenverteilung in den letzten Berichtsjahren. Im Berichtsjahr 2020 ist allerdings ein Wandel zu einer aktiveren Beteiligung der FDZ in der Sicherung der Datenqualität erkennbar, die nicht ausschließlich durch die neu akkreditierten FDZ zu begründen ist. Während alle fünf neuen FDZ sich grundsätzlich an der Qualitätssicherung, entweder teilweise oder ausschließlich, beteiligen, kümmern sich bei den länger bestehenden FDZ im Vergleich zum Vorjahr zwei FDZ nun ebenfalls zum Teil um die Sicherung der Datenqualität und zwei weitere FDZ haben diese Aufgabe ausgebaut.

Insgesamt ergibt sich somit für das Berichtsjahr 2020 folgendes Bild (► **Abb. 8**): Acht FDZ obliegt die Datenprüfung allein, bei 28 FDZ und somit über 70 % der FDZ erfolgt die Prüfung auf Qualität und Güte der Daten zumindest teilweise. In zwei FDZ werden keine Datenprüfungen vorgenommen.

Abb. 8: Datenprüfung in den FDZ

Gehört die Datenprüfung (auf Qualität und Güte der weitergegebenen Daten) zur Aufgabe Ihres FDZ?



Die Mehrheit der FDZ ist in die Datenprüfung eingebunden

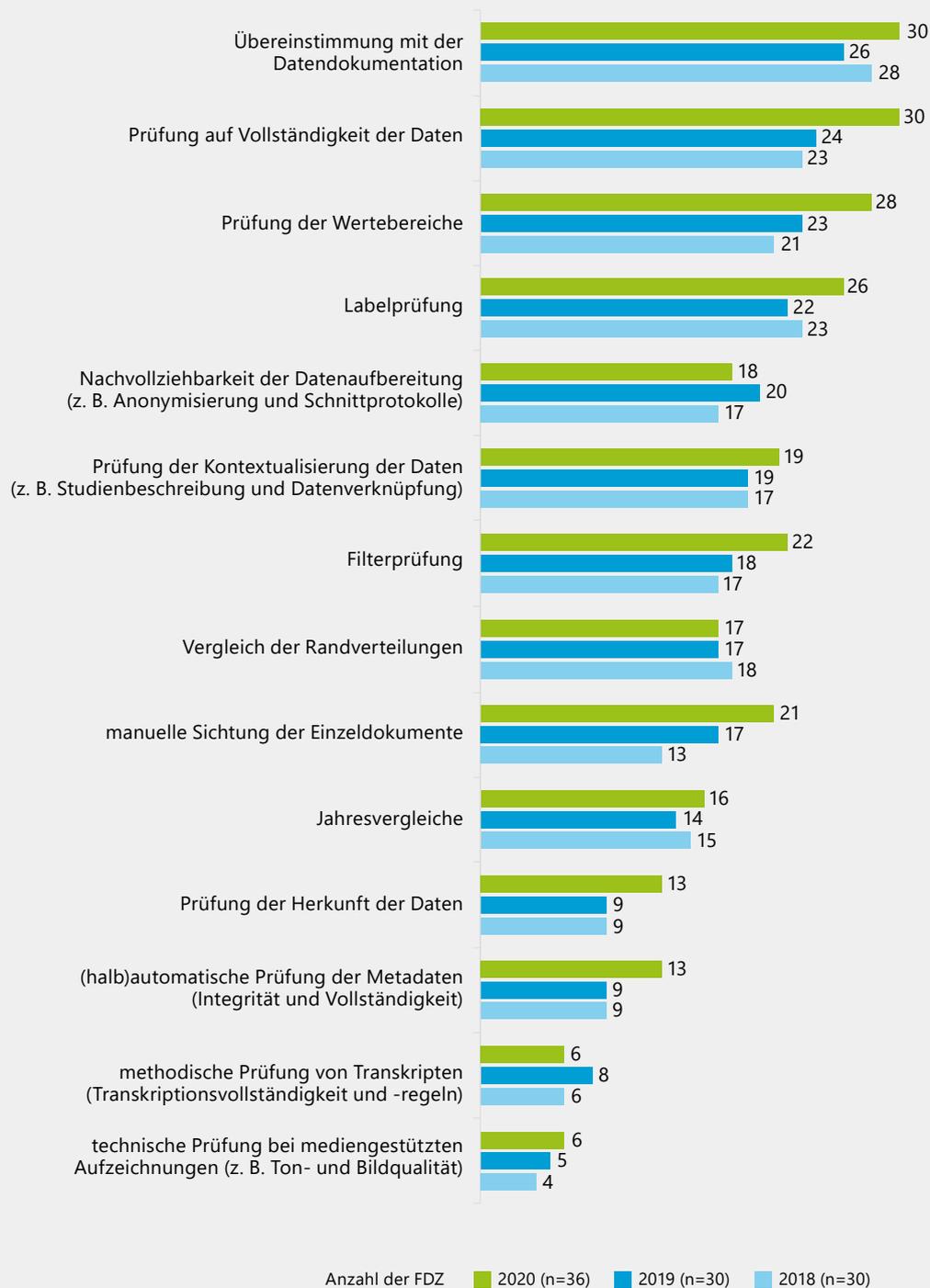
Am häufigsten prüfen die FDZ die Daten hinsichtlich der Übereinstimmung mit den Datendokumentationen, auf ihre Vollständigkeit und auf die Korrektheit der Wertebereiche und Labels (► **Abb. 9**). Die Anzahl von FDZ, die angeben die Herkunft der Daten zu kontrollieren oder auch einen automatischen Abgleich der Metadaten vorzunehmen, ist hingegen deutlich geringer. Ein Grund dafür kann sein, dass die Herkunft der Daten beispielsweise bereits vor der Übergabe an das FDZ geprüft wurde. Wenn bestimmte Datenprüfungen nicht im FDZ vorgenommen werden, heißt das nicht, dass sie nicht stattfinden, sondern an anderer Stelle im Lebenszyklus der Daten verankert sind.

¹⁰ Diese Neugestaltung ist dahingehend motiviert, dass die FDZ in den betreffenden Arbeitsbereichen etablierte, stabile Strukturen aufweisen.

Die methodische Prüfung von Transkripten und die technische Prüfung mediengestützter Aufzeichnungen werden in wenigen FDZ durchgeführt. Dies ist vor allem dadurch begründet, dass diese Arbeitsschritte in der Regel in der qualitativen empirischen Forschung zum Einsatz kommen, worauf nur ein kleiner Teil der akkreditierten FDZ spezialisiert ist.

Abb. 9: Art der Datenprüfung

Welche Art der Datenprüfung wird in Ihrem FDZ durchgeführt?
(Mehrfachantworten möglich)



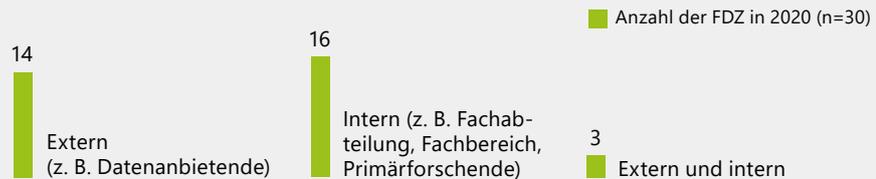
Die FDZ führen für die jeweiligen Datensätze passende Datenprüfungen durch; diese sind insbesondere abhängig von der Erhebungsmethode

Von denjenigen FDZ, denen die Datenprüfung nicht allein obliegt, werden am häufigsten die Primärforschenden oder Datenproduzierenden (dies können auch Drittinstitute sein) als weitere Akteur:innen bei der Qualitätssicherung benannt (► **Abb. 10**). In 23 FDZ gibt es für die Datenprüfung zudem ein Regelwerk.

Abb. 10: Von wem (noch) und wie wird die Datenprüfung durchgeführt?

(Auswertung offene Frage, Mehrfachnennungen möglich)

Die Datenprüfung erfolgt extern und intern



© RatSWD 2022

Die Datenprüfung ist ein wichtiger Schritt in der Qualitätssicherung der Daten. Ein weiterer Bestandteil ist die Korrektur von Datenfehlern einschließlich der Generierung von Daten, um beispielsweise die Nutzungsfreundlichkeit zu erhöhen. Hier ist die Aufgabenverteilung ähnlich wie bei der Datenprüfung: In drei Fällen obliegt die Korrektur ausschließlich, in 28 Fällen teilweise den FDZ. In sieben FDZ gehört die Datenkorrektur nicht zum Tätigkeitsspektrum.

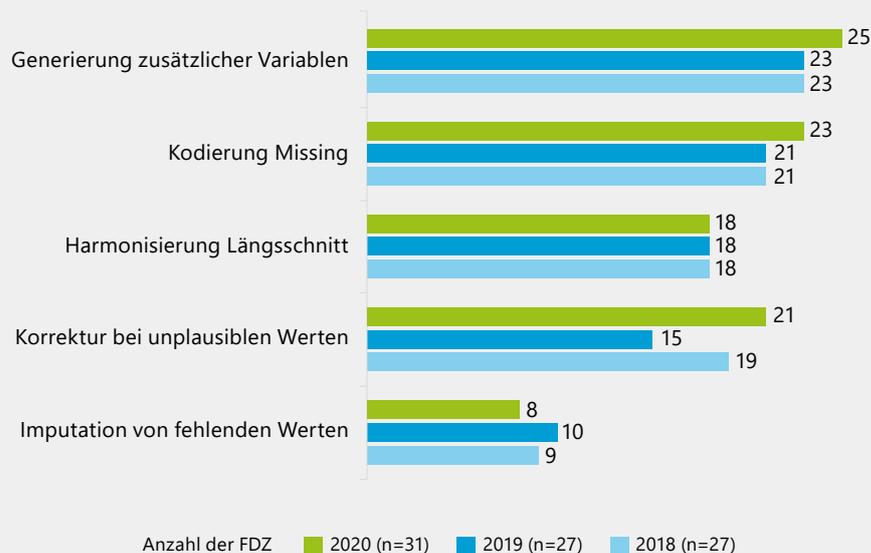
Zu den drei häufigsten Aufgaben in diesem Bereich gehörten in 2020 die Generierung zusätzlicher Variablen, die Kenntlichmachung fehlender Werte durch entsprechende Kodierungen sowie die Korrektur unplausibler Werte. Die Längsschnittharmonisierung gehört bei 18 FDZ zu den qualitätssichernden Korrekturen (► **Abb. 11**). In 18 der 31 FDZ, die Datenkorrekturen durchführen, gibt es dafür ein festes Regelwerk. Für die Datennutzenden sind die vorgenommenen Korrekturen nicht immer auch auf der Einzelfallebene erkennbar, z. B. wenn die Korrekturen aus Gründen des Datenschutzes oder der Anonymisierung erfolgten.

Abb. 11: Art der Datenkorrektur in den FDZ

Welche Art der Generierung oder Datenkorrektur wird in Ihrem FDZ durchgeführt?

(Mehrfachantworten möglich)

Datenkorrekturen erhöhen die Qualität und Nutzbarkeit der Daten



© RatSWD 2022

4

Datenangebot und Datennutzung

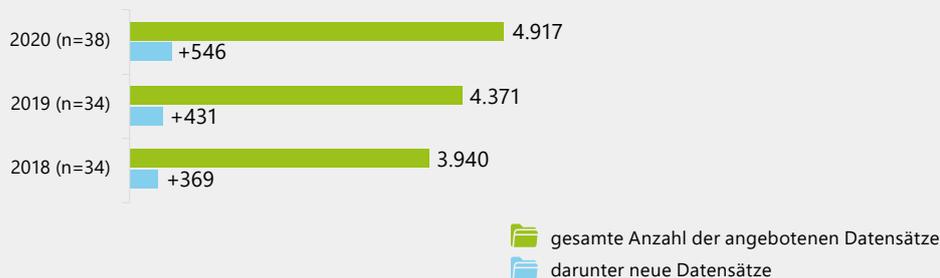
In diesem Kapitel wird das Datenangebot der FDZ, die Datennutzung und der Forschungsoutput näher betrachtet. Der letzte Abschnitt stellt differenziert dar, wie häufig Daten zum freien Download abgerufen wurden, wie viele Datennutzungsverträge im Jahr 2020 neu abgeschlossen wurden und wie viele Datennutzende neu hinzugekommen sind.

Umfang des Datenangebotes

Der Zuwachs im Datenangebot der letzten Jahre setzt sich fort: Zum Stichtag 31.12.2020 haben die FDZ 4.917 Datensätze angeboten. Damit sind im Berichtsjahr 2020 546 Datensätze neu hinzugekommen, die entweder einen neuen Digital Object Identifier (DOI) erhalten haben oder prinzipiell dazu geeignet gewesen wären (► **Abb. 12**). Hierbei variiert in den FDZ die Anzahl der neu angebotenen Datensätze, beginnend bei einer einstelligen Anzahl bis zu 130 neuen Datensätzen in einem FDZ. Ein Datensatz kann aus mehreren Einzelstudien bestehen, sodass die Anzahl bereitgestellter Studien wesentlich höher liegt.

Abb. 12: Angebotene Datensätze in den FDZ

Bitte geben Sie die Anzahl der im Berichtsjahr 2020 hinzugekommenen Erhebungen und Datensätze an. Als neu gelten Datensätze, die z. B. eine neue DOI erhalten haben oder prinzipiell dazu geeignet wären.



Die Anzahl der angebotenen Datensätze stieg in den vergangenen Jahren an

© RatSWD 2022

Für die angebotenen Datensätze vergeben 33 FDZ Persistente Identifikatoren (PID), wie z. B. DOI oder Uniform Resource Names (URN), um deren Auffindbarkeit und Zitierbarkeit dauerhaft zu gewährleisten. Alle FDZ, die noch keine PID vergeben, planen aktuell die Einführung eines solchen Verfahrens.

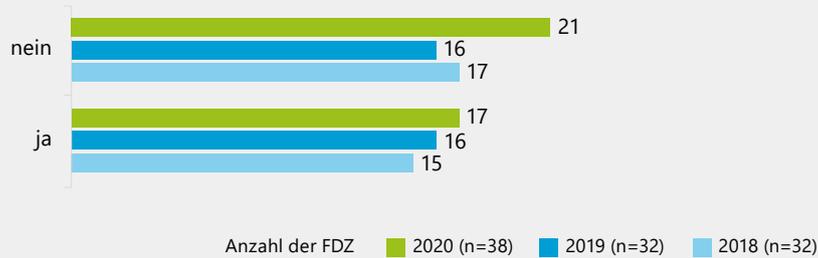
Zeitpunkt der Verfügbarmachung der Datensätze und Gebühren

Das zentrale Ziel der FDZ ist ein niedrigschwelliger und zeitnaher Datenzugang. Um die Chancengerechtigkeit zu wahren, sollten die Forschungsdaten allen qualifizierten Nutzenden zeitgleich zugänglich gemacht werden. Hohe Gebühren und lange Sperrfristen zur Datenweitergabe können diesen Zielen entgegenstehen. Gleichzeitig können aber auch gute Gründe vorliegen, warum Sperrfristen für die Datenweitergabe verhängt werden. So begründen die meisten FDZ Sperrfristen mit Erstverwertungsmöglichkeiten und -rechten der Primärforschenden. Damit werden etwa laufende Qualifikationsarbeiten durch Sperrfristen geschützt.

Über die Hälfte der FDZ bieten Datensätze ohne Sperrfristen an

Abb. 13: Sperrfristen für Datensätze

Gibt es Sperrfristen für einige Datensätze in Ihrem FDZ?



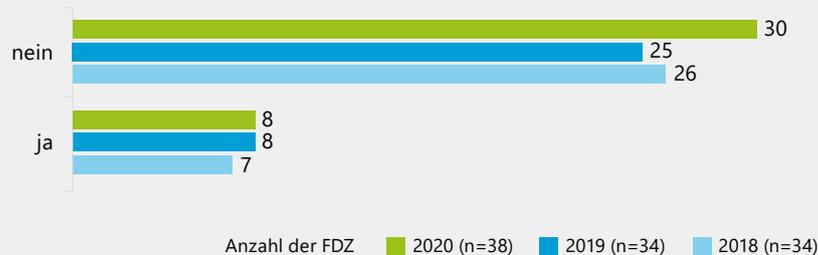
© RatSWD 2022

In 21 der 38 FDZ existieren keine Sperrfristen, d. h. die Daten werden unmittelbar nach der Aufnahme und Aufbereitung zur Verfügung gestellt (► **Abb. 13**). 17 FDZ gaben für das Berichtsjahr 2020 an, dass sie zumindest einen Teil ihrer Datenbestände mit Sperrfristen versehen. Fünf FDZ nennen feste Sperrfristen von sechs Monaten bis maximal zwei Jahren. Bei den übrigen FDZ sind die Sperrfristen nicht grundsätzlich festgelegt, sondern abhängig von bestimmten Bedingungen, wie beispielsweise dem Laufzeitende von Forschungsprojekten oder Vorgaben fördernder Institutionen bzw. der Datengebenden selbst. Insgesamt ist über die letzten Jahre eine gewisse Dynamik in der Nutzung von Sperrfristen zu bemerken; die absolute Anzahl der FDZ, die Datensätze mit einer Sperrfrist versehen, ist aufgrund von Neuakkreditierungen geringfügig gestiegen.

Die große Mehrheit der FDZ erhebt keine Gebühren bei der Datenbereitstellung

Abb. 14: Gebühren der Datenbereitstellung

Erhebt Ihr FDZ Gebühren bei der Datenbereitstellung für die wissenschaftliche Nutzung?



© RatSWD 2022

Neben den Sperrfristen kann eine Nutzungsgebühr ein Hemmnis für die Datennutzung sein. Ein Großteil der FDZ verzichtet völlig auf eine solche Gebühr: Von 38 FDZ gaben 30 für das Jahr 2020 an, dass sie keine Gebühren erhoben haben (► **Abb. 14**). Verglichen mit dem Vorjahr zeigt sich hier kaum Veränderung. Die genannten Gebühren der acht FDZ liegen zum größten Teil im zweistelligen bzw. unteren dreistelligen Euro-Bereich, häufig mit Ermäßigungen für die Durchführung von Qualifikationsarbeiten.¹¹ Wofür die Gebühren erhoben werden, differiert dabei zwischen den FDZ: Teilweise werden die Gebühren pro Datensatz oder Zugangsweg erhoben, teilweise wird die Datennutzung pro Erhebungsjahr berechnet. Auch der Aufwand für Datenaufbereitungen, beispielsweise für eine speziell notwendige Anonymisierung, oder für Beratungsleistungen wird teilweise in Rechnung gestellt. Gerade die niedrigen Beträge dürften allerdings nur wenig zur Kostendeckung des Aufwands in den FDZ beitragen. Sie dienen dann jedoch im Sinne einer Schutzgebühr auch dazu, die Ernsthaftigkeit der Datennutzung sicherzustellen.

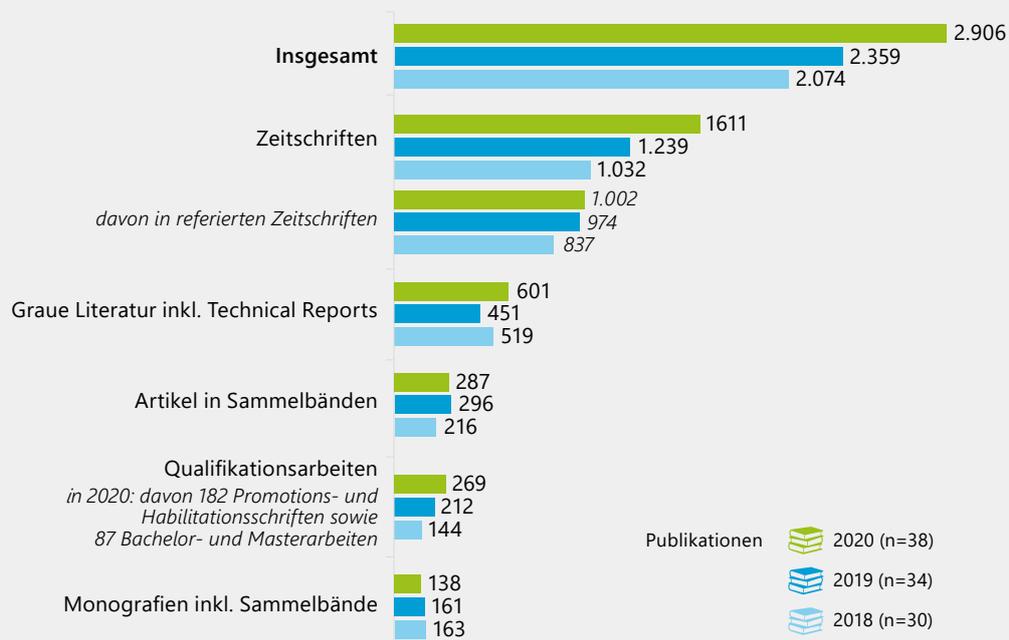
¹¹ Seit Juni 2020 erstattet das Statistische Bundesamt bei Qualifikationsarbeiten die bereits vergünstigten Kosten bei der Datennutzung für Nachwuchswissenschaftler:innen komplett. Somit kann der Datenzugang für Bachelor-, Master- und Promotionsarbeiten im Statistischen Verbund kostenlos angeboten werden.

Forschungoutput auf Basis der bereitgestellten Forschungsdaten

Die Nutzung der bereitgestellten Datensätze durch die Forschungsgemeinschaft zeigt sich zentral im Forschungoutput. Die Anzahl der Publikationen ist ein wesentlicher Indikator für diese Messgröße. Im Jahr 2020 meldeten die 38 FDZ insgesamt 2.906 Publikationen (► **Abb. 15**), die auf Basis der bereitgestellten Forschungsdaten erschienen sind. Dabei ist zu berücksichtigen, dass Forschende häufig die Rückmeldung oder Abgabe von Belegexemplaren an die FDZ versäumen. Obwohl die FDZ mit der breiten Nutzung von PID für Forschungsdaten die Möglichkeit zur Datenzitation gelegt haben, berücksichtigen nicht alle Autor:innen diese Zitationsmöglichkeiten in ihren Publikationen. Deshalb ist von einer erheblichen Untererfassung auszugehen.

Abb. 15: Publikationen auf Basis der bereitgestellten Forschungsdaten in den FDZ

Bitte geben Sie die Anzahl der im Berichtsjahr 2020 erschienenen wissenschaftlichen Publikationen an, die auf Basis der bereitgestellten Forschungsdaten erschienen sind.



Die Anzahl der Publikationen steigt kontinuierlich

Auch wenn die Ermittlung der Anzahl der auf der Grundlage der bereitgestellten Daten verfassten Publikationen unvollständig ist, kann im langjährigen Vergleich insgesamt ein Anstieg der erfassten Publikationen verzeichnet werden. Artikel in Fachzeitschriften stellen auch im Jahr 2020 die häufigste von den Datennutzenden gemeldete Publikationsart dar. Dabei müssen qualitative Unterschiede zwischen den Publikationsformen unterschieden werden. So sind Artikel in referierten Zeitschriften als höherwertig zu betrachten als in Fachzeitschriften ohne Peer-Review-Verfahren. Der Anteil in referierten Fachzeitschriften ist im Jahr 2020 besonders hoch.

Aber auch die Anzahl grauer Literatur einschließlich technischer Reports und der Qualifikationsarbeiten, die mit den Daten aus den FDZ erstellt werden, hat deutlich zugenommen.¹² Die Anzahl der Artikel in Sammelbänden hat geringfügig, die der Monographien sogar deutlich abgenommen.

¹² Bei den Qualifikationsarbeiten muss ebenfalls von einer erheblichen Untererfassung ausgegangen werden, da insbesondere Bachelor- und Masterarbeiten im Allgemeinen nur aufgrund der Rückmeldungen der Datennutzenden erfasst werden können und sie über Recherche schwer auffindbar sind. Hinzu kommt, dass veröffentlichte Promotionsarbeiten teilweise innerhalb größerer Projekte erstellt werden und die Datennutzung der Autor:innen nicht in jedem Fall bei den FDZ registriert wird. Es ist also davon auszugehen, dass noch wesentlich mehr Qualifikationsarbeiten unter Rückgriff auf FDZ Daten verfasst werden.

Nutzungsumfang

Der Nutzungsumfang stellt einen weiteren zentralen Indikator für die Bedeutung der FDZ in der Forschungslandschaft dar. Ein zentraler Wert ist dabei die Anzahl der über die FDZ bezogenen Datensätze. Daneben interessiert auch, wie viele Forschende von den Leistungen der FDZ profitieren konnten.

Da einige FDZ stark an die eigene Institution angebunden sind, andere eher eine Dienstleistungsinfrastruktur mit einer starken Außenorientierung aufweisen, ist es nicht ganz trivial, die Nutzungszahlen einzuordnen.

Zu den zentralen Aufgaben der FDZ gehört es, für die Forschung einen umfangreichen und flexiblen Datenzugang zu schaffen und diesen ständig auszubauen und zu verbessern. Für Forschende werden unter anderem aufgrund unterschiedlicher Datenschutzerfordernungen und rechtlicher Regelungen verschiedene Zugangsmöglichkeiten angeboten. Aufgrund dieser Flexibilität von Zugangsmöglichkeiten und der unterschiedlichen Erfassung von Verträgen, Projekten und Datennutzenden in den einzelnen FDZ ist es schwierig, eine präzise Zahl für den kompletten Nutzungsumfang zu ermitteln. So konnten in den Vorjahren zum Teil Doppelzählungen bei Verträgen und Projekten oder zwischen den einzelnen Zugangswegen nicht ausgeschlossen werden.

Eine solche Vielfalt empirisch abzubilden, erweist sich als komplex. Deshalb werden seit dem Tätigkeitsbericht 2019 vier Teilindikatoren differenziert: Anzahl der freien Downloads, der externen Datennutzenden, der Datennutzungsverträge und der davon umfassten Nutzenden bzw. Datensätze.



68.752

Freie Downloads
n=17

Downloads

Insgesamt ist es an 17 FDZ möglich, Datensätze frei per Download zu beziehen. Bei einem Großteil der FDZ erfordert der freie Download eine Registrierung. Einige frei verfügbare Datensätze können aber auch ohne vorhergehende Registrierung heruntergeladen werden – was die Identifikation individueller Nutzender erschwert. Die Zählung der Nutzenden kann daher nicht für alle Datenzugriffe dokumentiert werden. Teilweise kann auch aufgrund der technischen Gegebenheiten die genaue Anzahl der Downloads nicht ermittelt werden. Erschwerend kam hinzu, dass bei einem großen FDZ aufgrund technischer Probleme zu Beginn des Jahres 2020 der Datenzugang zunächst nicht möglich war und in der Folgezeit die Anzahl an Downloads aufgrund technischer Umstellungen zudem nicht genau ermittelt werden konnte. Insgesamt wird deshalb die Anzahl der Downloads im Jahr 2020 noch stärker unterschätzt als in den Vorjahren. Die zwölf FDZ, die unmittelbar zu den Downloadzahlen Angaben machen können, geben an, dass im Berichtsjahr 2020 frei verfügbare Datensätze 68.752-mal heruntergeladen wurden.

In der Regel ist das Forschungsdatenangebot der FDZ erst nach Vertragsabschluss oder Registrierung zugänglich.



43.703

Externe Datennutzende
n=28

+ 7.662

Neue externe
Datennutzende
n=28

Externe Datennutzende

Insgesamt wurden im Berichtsjahr 43.703 externe Datennutzende der FDZ gezählt. Als externe Datennutzende werden in der Regel Datennutzende verstanden, die nicht dem jeweiligen FDZ bzw. der Institution, zu der das FDZ zählt, angehören. 28 FDZ konnten Angaben zur Anzahl der externen Datennutzenden machen.

Im Jahr 2020 hat die Anzahl der externen Datennutzenden um 7.662 Personen zugenommen. Angaben zur Anzahl der im Jahr 2020 neu hinzugekommenen Datennutzenden konnten 28 FDZ machen. Zwei FDZ hatten im Jahr 2020 technische Schwierigkeiten in der Erfassung der neuen Datennutzenden. Das bedeutet, dass die Anzahl der in 2020 neu hinzugekommenen Datennutzenden insgesamt noch höher liegen dürfte und hier eher untererfasst wird.

Datennutzungsverträge

Aus Datenschutzgründen enthalten Datenbereitstellungs- und Nutzungsverträge einen expliziten Bezug zu einem Forschungsprojekt oder Qualifikationsvorhaben (Zweckbindung), d. h. für jedes Forschungsprojekt, das solche Daten nutzt, muss ein Vertrag geschlossen werden. Für die Gestaltung des Vertrages gibt es jedoch keine formalen Vorlagen. Die konkrete Vertragsgestaltung unterliegt der Vertragsfreiheit, die Regelungstiefe ist abhängig von gesetzlichen Vorgaben und Erfordernissen. So ist die gesetzliche Reglementierung bei Daten der amtlichen Statistik in der Regel recht umfassend durch Verordnungen oder Gesetze geregelt. Für viele Befragungsdaten müssen vor allem Regelungen des Datenschutzes eingehalten werden. Dies gilt insbesondere für sensible personenbezogene Informationen. Andere Daten, etwa zur Regionalisierung oder Flächennutzung, sind für bestimmte Nutzungsfälle frei verfügbar oder unterliegen bestimmten Lizenzen. Dementsprechend vielfältig zeigt sich die Vertragsgestaltung. Dies gilt nicht nur für die Daten an sich, sondern auch für die vertragsschließenden Parteien: Verträge werden mit Einzelpersonen, Projekten oder Instituten geschlossen. Teilweise umfassen sie ganze Datenbestände, Studienkollektionen oder nur einzelne Datensätze.

Im Jahr 2020 gab es an 37 FDZ 38.219 laufende Datennutzungsverträge in der Forschungsdateninfrastruktur des RatSWD. An 33 FDZ wurden 6.321 neue Datennutzungsverträge im Jahr 2020 abgeschlossen.

14 FDZ gaben an, dass für jede datennutzende Person ein eigener Datennutzungsvertrag abgeschlossen werden muss. 24 FDZ schließen einen Vertrag für mehrere Nutzende ab. 24 FDZ regeln mit einem Datennutzungsvertrag den Zugriff auf mehrere Datensätze (► **Abb. 16**).



38.219

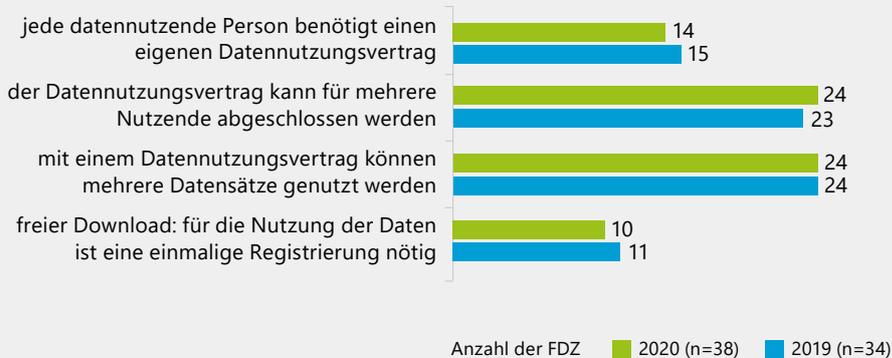
Laufende
Datennutzungsverträge
n=37

+ 6.321

Neu abgeschlossene
Datennutzungsverträge
im Jahr 2020
n=33

Abb. 16: Vertragsgestaltung

Was trifft auf Ihr FDZ zu? (Mehrfachantworten möglich)



Die Vertragsgestaltung variiert stark: Verträge können eine oder mehrere Personen oder Datensätze umfassen

Im Falle eines Vertragsabschlusses mit mehreren Nutzenden erlauben 19 FDZ die Nutzung ausschließlich durch spezifisch benannte Personen. Im Durchschnitt erhielten dabei 2,8 Personen über einen Vertrag Zugriffsrechte. 16 FDZ schließen projektbezogene Verträge und zehn FDZ erlauben Datenzugriffe für gesamte Institute.

Wenn im Rahmen eines Vertrags auf mehrere Datensätze zugegriffen werden kann, darf in zwölf FDZ eine benannte Person diese nutzen, in 15 FDZ mehrere Personen. Durchschnittlich wurden dann etwas mehr als zwei Personen je Vertrag genannt. Eine projektbezogene Nutzung erlauben neun FDZ. Eine Ausnahme stellt eine institutsweite Nutzung dar, die lediglich ein FDZ berichtet.

5 Etablierte Datenzugänge und Servicekonzepte

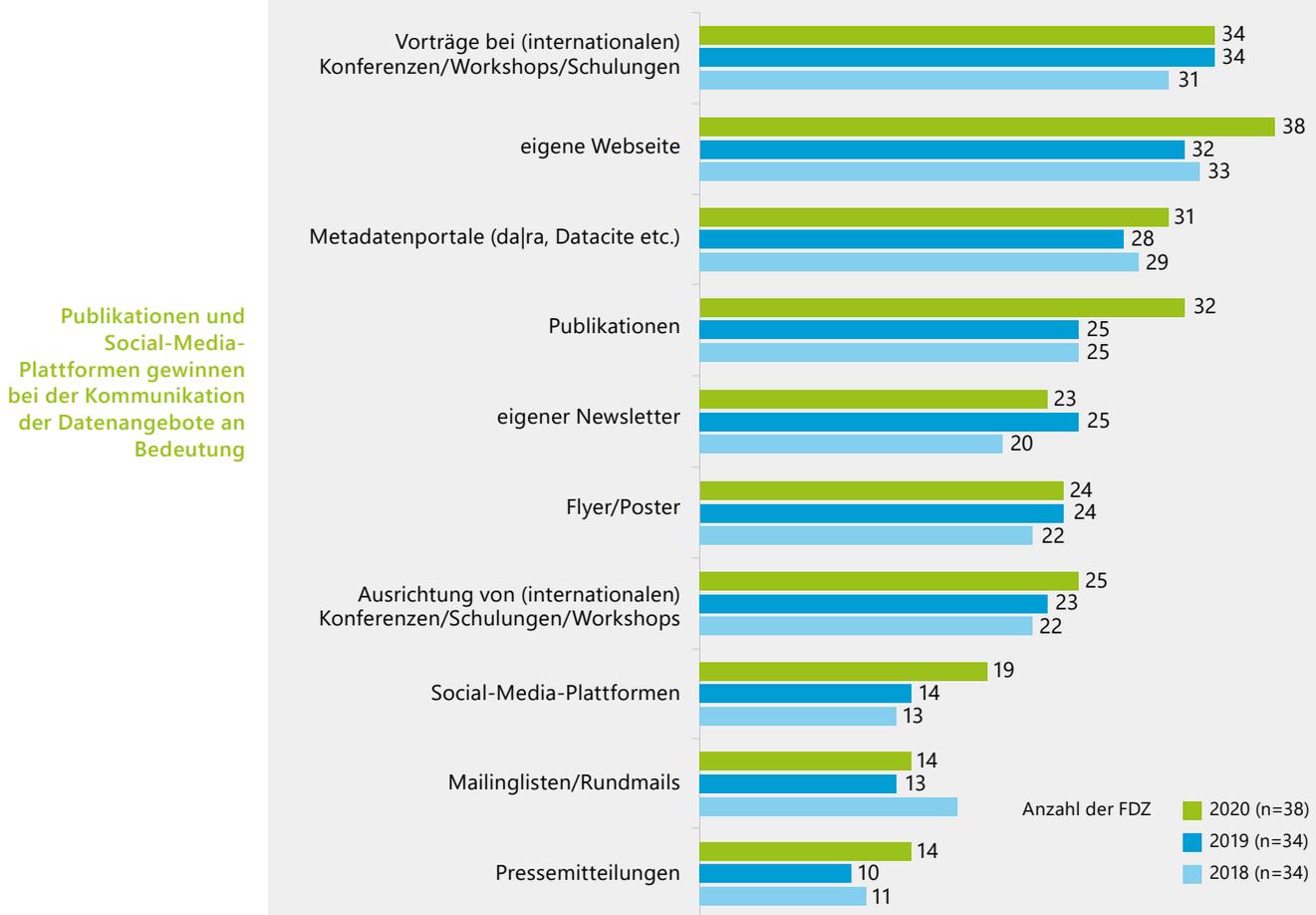
In diesem Kapitel werden die in den FDZ etablierten Datenzugänge sowie das breite Spektrum an Servicekonzepten zur Unterstützung der Datennutzenden im Einzelnen näher betrachtet. Der erste Abschnitt stellt zudem die verschiedenen Instrumente vor, die in den FDZ zur Bewerbung des eigenen Datenangebots verwendet werden.

Bewerbung des Datenangebots

Die FDZ verwenden eine Bandbreite an Kommunikationskanälen, um die wissenschaftliche Gemeinschaft über das eigene Datenangebot und dessen Potenziale zur Beantwortung wissenschaftlicher Fragestellungen zu informieren. Vorherrschende Instrumente der Eigenwerbung sind die eigene Webseite (n=38), Vorträge auf (inter-)nationalen Konferenzen, Schulungen und Workshops (n=34) sowie die Verwendung von Metadatenportalen wie da|ra oder Datacite (n=31). Zudem ist im Berichtsjahr 2020 bei Publikationen als

Abb. 17: Genutzte Informationskanäle zur Kommunikation des Datenangebots durch die FDZ

Wie und über welche Wege werden die angebotenen Daten der Wissenschaft bekannt gemacht? (Mehrfachantworten möglich)



Werbeinstrument ein Anstieg von 25 FDZ im Vorjahr auf nun 32 FDZ zu verzeichnen. Ebenso ist ein Anstieg in der Akzeptanz von Social-Media-Plattformen zur Eigenwerbung erkennbar. Während im Vorjahr noch 14 FDZ angaben, Social-Media-Plattformen zur Bewerbung des eigenen Datenangebots zu verwenden, sind dies im Berichtsjahr 2020 nun 19 FDZ (► **Abb. 17**). Hierbei ist anzumerken, dass die Anstiege in Publikationen und Social-Media-Plattformen als Kommunikationskanäle nicht nur ausschließlich durch die neu akkreditierten FDZ zu erklären sind, sondern auch einige „erfahrene“ FDZ diese Kanäle für sich entdeckt haben.

Zugangswege und -formate für die Datenbereitstellung

Wie schon in Kapitel 4 ausgeführt, boten die FDZ zum Stichtag 31.12.2020 insgesamt 4.917 Datensätze an. All diese Mikrodaten unterscheiden sich selbstverständlich nicht nur thematisch, sondern auch in ihrem Anonymisierungsgrad, wobei letzteres die möglichen Datenzugangswege entscheidend bestimmt.

Infokasten 2

Datenzugangswege

Gastwissenschaftsarbeitsplätze (GWAP)

Die Daten werden den Forschenden an speziell gesicherten Arbeitsplätzen in den FDZ bereitgestellt. Typische Merkmale dieser GWAP sind zum Beispiel, dass sie keinen unkontrollierten Netzzugang besitzen und ein lokales Abspeichern der Daten nicht möglich ist. Die Ergebnisdateien werden vor der Übermittlung auf die Einhaltung des Datenschutzes geprüft (Outputkontrolle).

Kontrollierte Datenfernverarbeitung

Die Forschenden können die Daten per Fernzugriff auswerten, ohne selbst vor Ort im FDZ sein zu müssen. Die Datenspeicherung und -verarbeitung erfolgt weiterhin ausschließlich auf den Servern der FDZ und die Ergebnisdateien werden auch hier erst nach einer Prüfung durch das FDZ (Outputkontrolle) an die Forschenden übermittelt. Grundsätzlich können hierbei zwei Verfahren unterschieden werden:

■ Remote Execution:

Datenzugangsweg, bei dem die Forschenden die Daten und die Zwischenergebnisse nicht unmittelbar einsehen können. Die Forschenden schreiben die Analyseskripte an ihren eigenen Arbeitsplätzen (eventuell mit Hilfe eines Strukturdatensatzes, an dem sie die Programmsyntax testen können) und übermitteln diese an das FDZ, das diese auf die Originaldaten anwendet.

■ Remote Desktop:

Datenzugangsweg, bei dem die Benutzendenoberfläche des FDZ-Servers mit Hilfe eines Remote Access' auf den Bildschirm des lokalen Clients der Forschenden übertragen wird. Über eine geeignete Konfiguration kann verhindert werden, dass die Daten auf dem lokalen Client gespeichert werden können. Das lokale Zugangsgerät dient lediglich dazu, mit dem Datenserver zu kommunizieren. Forschende nutzen auf dem Server die Analysesoftware und arbeiten mit den Forschungsdaten als würden sie lokal vorliegen.

Bei weiteren Datenzugangsweegen werden die Daten den Nutzenden per Download, digitalem oder postalischen Versand bereitgestellt und können von ihnen unmittelbar auf ihrem lokalen Computer analysiert werden. Die Zugangswege unterscheiden sich in ihrem Anonymisierungsgrad und Nutzungszweck:

Scientific Use Files (SUF)

Für wissenschaftliche Forschungszwecke bereitgestellte Datensätze, die grundsätzlich mindestens faktisch anonymisiert sind, aber dennoch ein hohes Analysepotenzial bieten.

Campus Files (CF)

Für die universitäre Lehre bestimmte und im Vergleich zu Scientific Use Files noch stärker anonymisierte Datensätze.

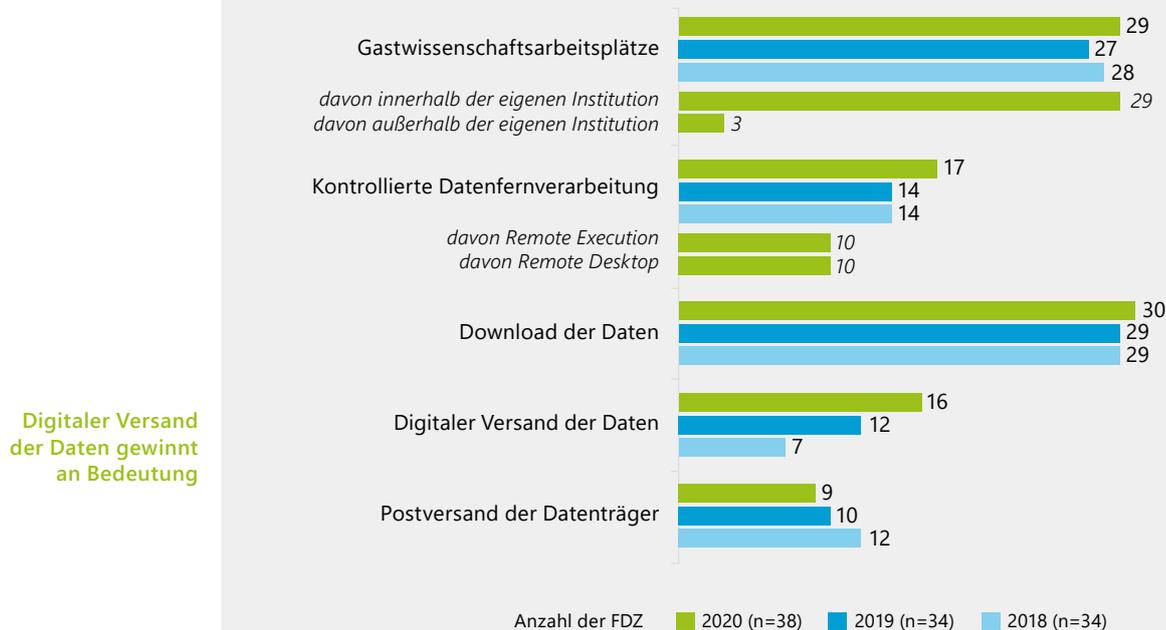
Public Use Files (PUF)

Anonyme Datensätze, bei denen keine Nutzungseinschränkungen bestehen und somit eine Weitergabe der Daten auch außerhalb der wissenschaftlichen Forschung möglich ist.

In der EU-Datenschutz-Grundverordnung (EU-DSGVO) wurden die bisher verwendeten Begrifflichkeiten formale und faktische Anonymisierung abgeschafft.¹³ Es gibt nur noch die Unterscheidung zwischen Pseudonymisierung und Anonymisierung. Für eine bessere Unterscheidung der Anonymisierungsgrade der Datenzugangswege werden die alten Begrifflichkeiten im Text beibehalten. Formal anonymisierte Mikrodaten können nur unter hohen technisch-organisatorischen Maßnahmen zur Verfügung gestellt werden. Hierbei haben sich der GWAP wie auch die Datenfernverarbeitung als Datenzugangswege etabliert. Faktisch oder gar absolut anonymisierte Mikrodaten wie Scientific Use Files (SUF), Campus Files (CF; auch: Campus Use Files, CUF) oder Public Use Files (PUF) werden standardmäßig als Downloads bereitgestellt bzw. können zudem postalisch oder digital versendet werden. ► **Infokasten 2, S. 27** stellt die einzelnen Datenzugangswege im Detail vor.

Abb. 18: Angebotene Datenzugangswege

Über welche Zugangswege werden die Daten Ihres FDZ angeboten?
(Mehrfachantworten möglich)



© RatSWD 2022

Etwas mehr als drei Viertel (n=29) aller FDZ bieten Mikrodaten an einem GWAP an, wovon drei FDZ¹⁴ das eigene Datenangebot zudem an GWAP außerhalb der eigenen Institution zur Verfügung stellen. 17 FDZ ermöglichen einen Zugriff auf das eigene Datenangebot mit Hilfe einer Datenfernverarbeitung, entweder als Remote Execution oder Remote Desktop¹⁵ (► **Abb. 18**). Im Vorjahr boten nur 14 FDZ diesen Zugangsweg an. Dies entspricht nicht nur einem absoluten, sondern auch einem realen Zuwachs, da von den fünf neu akkreditierten FDZ derzeit nur eins einen Remote Desktop-Zugang anbietet.

Ebenfalls etwas mehr als drei Viertel (n=30) aller FDZ ermöglichen den geschützten Download von Mikrodaten. Leicht rückläufig ist das Angebot der postalischen Versendung physischer Datenträger, von zehn FDZ in 2019 auf neun FDZ in 2020. Der Trend geht eher zum digitalen Datenversand. Während im

¹³ In Erwägungsgrund 26 EU-DSGVO wird auf das Konzept der Anonymisierung Bezug genommen. Artikel 89 EU-DSGVO behandelt zudem das „Forschungsprivileg“ mit personenbezogenen Daten zu arbeiten, unter Einhaltung eines Anonymisierungsgebots für besondere Kategorien personenbezogener Daten.

¹⁴ Im Vorjahr wurden die FDZ des Bundes und der Länder getrennt voneinander befragt. Der Rückgang von vier FDZ im Berichtsjahr 2019 auf nun drei FDZ ist ausschließlich dadurch begründet, dass in diesem Berichtsjahr beide FDZ gemeinsam befragt wurden.

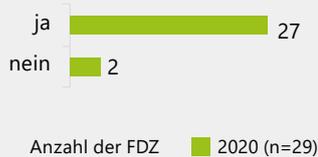
¹⁵ Einige FDZ verwenden Verfahren für die kontrollierte Datenfernverarbeitung, die keinen Einblick in die Daten, aber eine Vorabansicht auf die Ergebnisse ermöglichen.

Jahr 2018 sieben FDZ und im Jahr 2019 zwölf FDZ ihr Datenangebot über digitalen Datenversand bereitstellen, sind es im Berichtsjahr 2020 nun 16 FDZ. Dieser Anstieg ist nicht ausschließlich durch die neu akkreditierten FDZ begründet.

Die seit dem Berichtsjahr 2020 in Europa bestehende COVID-19-Pandemie, mit der teils strenge Ausgangs- und Kontaktbeschränkungen einhergingen, hatte auch einen negativen Einfluss auf den Betrieb der GWAP (► **Abb. 19**). 27 FDZ gaben an, dass aufgrund von Regelungen zur Eindämmung der COVID-19-Pandemie GWAP innerhalb der eigenen Institution zeitweise geschlossen wurden. Diese Einschränkung betraf auch alle FDZ, die GWAP außerhalb der eigenen Institution betreiben. Zudem ist anzumerken, dass auch bei Wiedereröffnungen von GWAP Einschränkungen bedingt durch die Regelungen des Infektionsschutzgesetzes, insbesondere in den Kapazitäten, zumindest teilweise vorherrschten.

Abb. 19: Auswirkungen der COVID-19-Pandemie 2020

19a: Mussten in Ihrem FDZ Gastwissenschafts-arbeitsplätze **innerhalb** der eigenen Institution aufgrund der COVID-19-Pandemie im Jahr 2020 geschlossen werden?



19b: Mussten in Ihrem FDZ Gastwissenschafts-arbeitsplätze **außerhalb** der eigenen Institution aufgrund der COVID-19-Pandemie im Jahr 2020 geschlossen werden?



Pandemiebedingte Schließung von Gastwissenschafts-arbeitsplätzen

Bearbeitungsdauer nach Eingang eines unterzeichneten Vertrags

Dieser Abschnitt betrachtet im Detail, wie viel Zeit ein FDZ durchschnittlich benötigt, um nach Eingang eines unterzeichneten Datennutzungsvertrags eine Freigabe der beantragten Mikrodaten bewilligen zu können. Hierbei geht es nicht um den tatsächlichen Zeitpunkt der Datenbereitstellung, da durch Datennutzende Verzögerungen im Zeitpunkt der tatsächlichen Bereitstellung entstehen können, die ein FDZ nicht zu verantworten hat. So können Datennutzende beispielsweise Mikrodaten, die als Download angeboten werden, erst zu einem späteren Zeitpunkt herunterladen, bei Mikrodaten, die an einem GWAP angeboten werden, hängt die tatsächliche Datenbereitstellung vom vereinbarten Termin des Gastaufenthaltes ab.

Die Dauer der Vorgangsbearbeitung in den FDZ hängt beispielsweise davon ab, ob Datensätze extra erstellt bzw. bestehende Datensätze zumindest angepasst werden müssen oder ob schon existierende Datensätze ohne eine notwendige Nachbearbeitung einem Datennutzenden bereitgestellt werden können. Zudem spielen auch verschiedene bürokratische Prozesse und deren Automatisierungsgrad innerhalb einer Institution eine Rolle.

Die Bearbeitungsdauer nach Eingang eines unterzeichneten Vertrags variiert zwischen den FDZ. Etwas mehr als die Hälfte (n=20) aller FDZ ist in der Lage, innerhalb einer Woche nach Eingang des unterzeichneten Vertrags die beantragten Datensätze zur Verfügung zu stellen. Etwas mehr als ein Viertel (n=11) aller FDZ kann gewöhnlich innerhalb von vier Wochen Datensätze bereitstellen, bei vier FDZ dauert eine Bereitstellung in der Regel länger als vier Wochen.

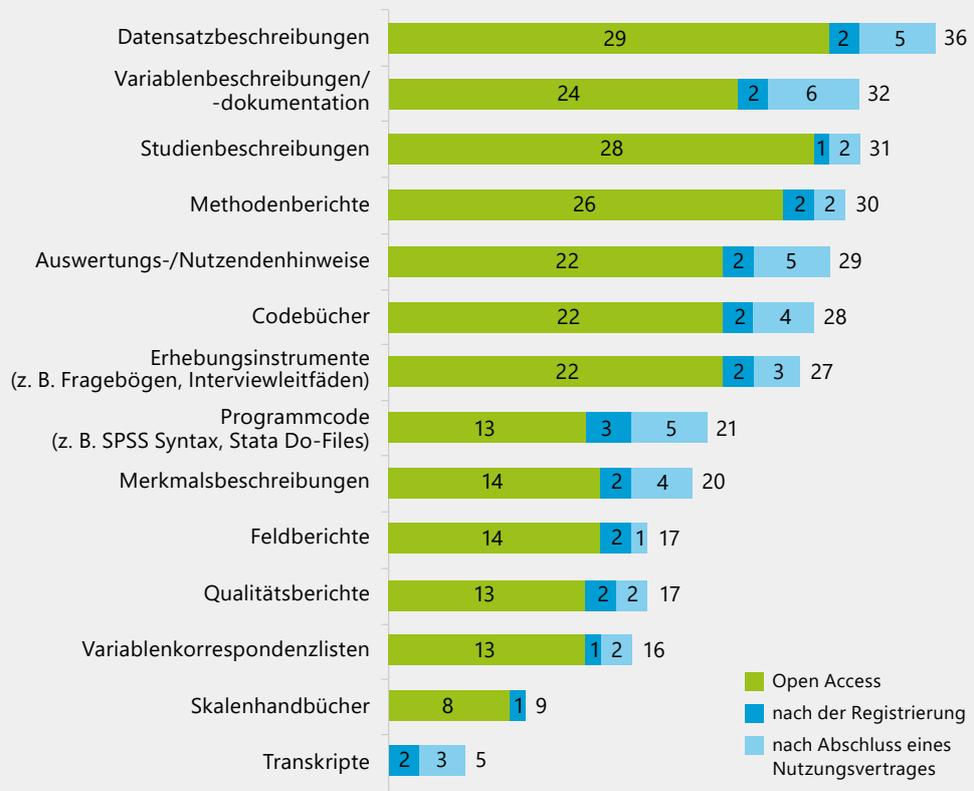
Bereitstellung von Arbeitshilfen

Datensatzspezifische Arbeitshilfen sind für Datennutzende hilfreich, um einen guten Einblick in das Potenzial eines Datensatzes für die Beantwortung bestimmter Fragestellungen zu gewinnen. Dabei ist eine frei zugängliche Bereitstellung datensatzspezifischer Arbeitshilfen für Datennutzende wie auch für die FDZ von Vorteil: Mit Hilfe von Datensatzbeschreibungen und Codebüchern können Datennutzende sich schon vor einer Antragsstellung besser über das Potenzial eines Datensatzes für ihre Forschungsprojekte informieren, unnötige Vertragsabschlüsse und somit auch Arbeitsaufwände werden vermieden. Die nachfolgende ► **Abb. 20** zeigt, welche Arbeitshilfen über welchen Zugangsweg bereitgestellt werden.

Abb. 20: Bereitstellung von Arbeitshilfen in 2020

Welche Arbeitshilfen werden über welchen Zugangsweg bereitgestellt? (n=38)
(Mehrfachantworten möglich)

Die FDZ stellen umfangreiche und vielfältige Arbeitsmaterialien zur Verfügung



© RatSWD 2022

Wie die Abbildung zeigt, stellen fast alle FDZ zu jedem angebotenen Datensatz zusätzliche Informationsmaterialien zur Verfügung. Diese werden mehrheitlich per Open Access¹⁶ über die Website eines FDZ angeboten, sodass Datennutzende sich bereits vor einer Antragsstellung über einen Datensatz informieren können. Unabhängig vom Zugangsweg werden von den FDZ überwiegend Datensatz-, Variablen- und Studienbeschreibungen sowie Methodenberichte als Arbeitshilfen bereitgestellt.

Darüber hinaus bieten einzelne FDZ zusätzlich spezifischere Arbeitshilfen an. Hierzu gehören beispielsweise im Bereich der qualitativen Forschung Zusatzinformationen zu Tafelbildern, Sitzplänen oder Interviewprotokollen. Aber auch neuere Kommunikationsformen wie Video-Tutorials sowie interaktive Metadatenportale werden genutzt, um Datennutzende zu unterstützen.

¹⁶ Ziel von Open Access ist es, wissenschaftliche Literatur und Materialien für alle Nutzenden frei zugänglich zu machen; kostenlos und möglichst frei von technischen und rechtlichen Barrieren.

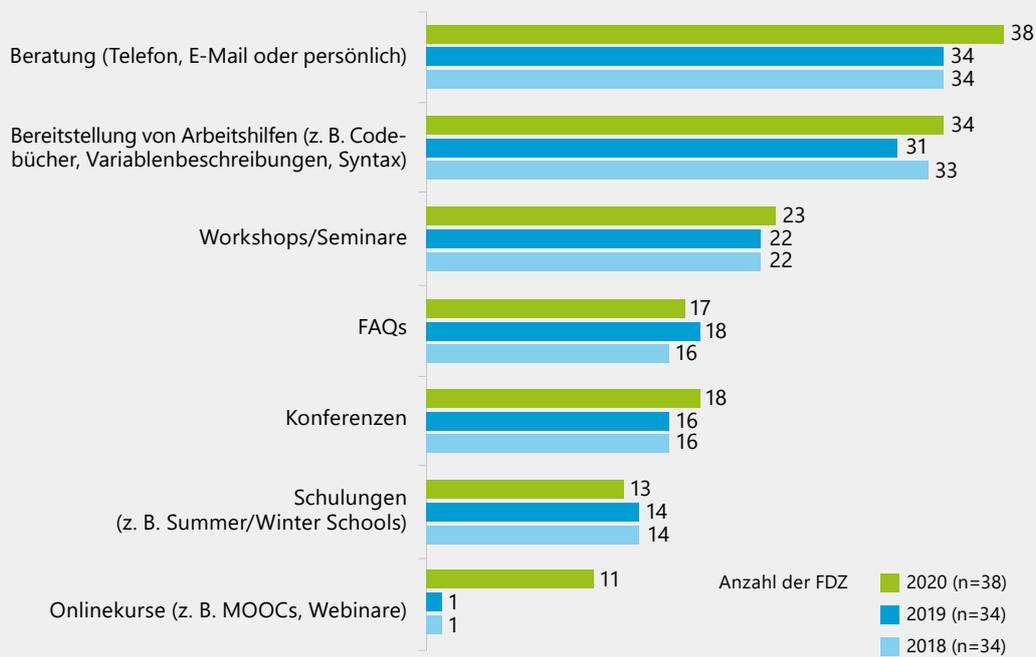
Serviceangebot für Datennutzende

Die Beratung und Unterstützung von Datennutzenden ist neben der Aufbereitung von Mikrodaten sowie der Erstellung von Arbeitshilfen eine wichtige Säule der Arbeit der FDZ. Wie in Kapitel 2 im Detail behandelt, kann das Personal eines FDZ einen Teil der eigenen Arbeitszeit in aktive Forschung investieren. Diese Erfahrung im Umgang mit dem eigenen Datenangebot zur Beantwortung wissenschaftlicher Fragestellungen ist insbesondere für die Beratung von externen Datennutzenden von enormem Vorteil, da das FDZ Personal hierdurch mögliche Fallstricke und Potenziale eines Datensatzes besser überschauen und vermitteln kann. Nachfolgend wird das Serviceangebot der FDZ näher dargestellt.

Abb. 21: Nutzendenservice in den FDZ

Welche Services werden für die Nutzenden durch Ihr FDZ angeboten?

(Mehrfachantworten möglich)



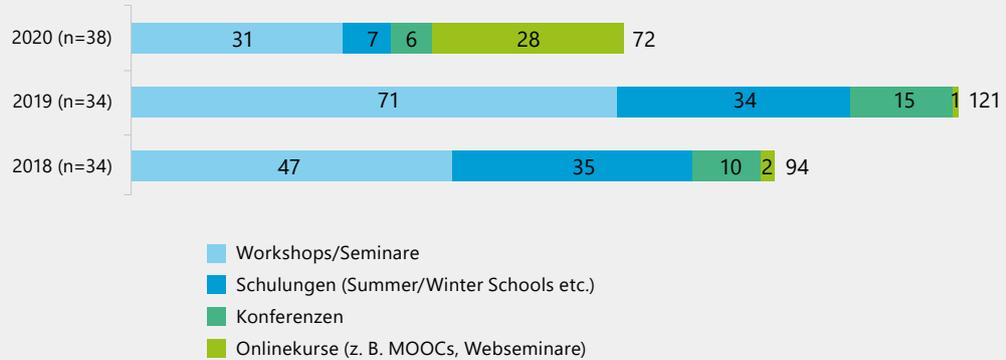
Alle FDZ fokussieren vor allem auf eine individuelle Nutzungsberatung

Die individuelle Beratung von Datennutzenden per Telefon, E-Mail oder auch vor Ort wird von allen FDZ angeboten. Die Anzahl an FDZ, die Workshops/Seminare, Konferenzen und Schulungen anbieten, ist nahezu konstant geblieben. Von besonderem Interesse ist der im Vergleich zum Vorjahr starke Zuwachs an FDZ, die Onlinekurse wie Massive Open Online Courses (MOOCs) und Webinare anbieten. Während in 2018 und 2019 jeweils nur ein FDZ Onlinekurse anbot, sind dies im aktuellen Berichtsjahr insgesamt elf FDZ. Zusätzlich bieten FDZ vereinzelt virtuelle Hilfen in Form von Video-Tutorials und Diskussionsforen an (► **Abb. 21**). Dieser außerordentliche Anstieg kann zum größten Teil der allgemeinen Verlagerung bestimmter Lebensbereiche in den digitalen Raum aufgrund der COVID-19-Pandemie und dem Bestreben der FDZ, pandemiebedingte Einschränkungen in bestimmten „Offline“-Serviceangeboten durch zusätzliche Online-Varianten zu kompensieren, zugeschrieben werden. Dieser Trend ist auch bei den abgehaltenen Qualifizierungsmaßnahmen im Berichtsjahr 2020 erkennbar (► **Abb. 22, S. 32**).

Zunahme der Onlinekurse mit Beginn der Pandemie

Abb. 22: Qualifizierungsmaßnahmen

Wie viele der folgenden Qualifizierungsmaßnahmen bietet Ihr FDZ an?
(Mehrfachantworten möglich)



© RatSWD 2022

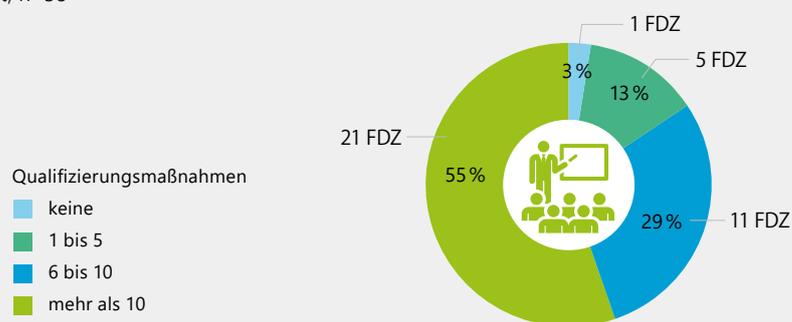
Während in 2019 und somit im letzten Jahr vor der COVID-19-Pandemie noch ein Anstieg auf 71 veranstaltete Workshops und Seminare beobachtbar ist, wurden im Berichtsjahr 2020 nur noch insgesamt 31 Workshops und Seminare abgehalten. Der gleiche negative Trend ist bei Schulungen, mit einer Abnahme von 34 Schulungen im Jahr 2019 auf sieben Schulungen im Jahr 2020, sowie bei Konferenzen, mit 15 Veranstaltungen im Jahr 2019 und sechs Veranstaltungen im Jahr 2020, zu beobachten. Die Organisation von Onlinekursen stieg dagegen stark an. Während im Jahr 2019 nur ein FDZ einen Onlinekurs veranstaltete, boten im Berichtsjahr 2020 neun FDZ insgesamt 28 Veranstaltungen an. Dieser Anstieg ist hierbei insbesondere drei FDZ geschuldet, die ihr Angebot an Onlinekursen stark ausgeweitet haben. Im Vergleich zum Vorjahr sank die Anzahl organisierter Qualifizierungsmaßnahmen von 121 auf insgesamt 72 Veranstaltungen.

Die im Berichtsjahr 2020 angebotenen Qualifizierungsmaßnahmen sind über die FDZ unterschiedlich verteilt (► **Abb. 23**). Mehr als die Hälfte aller FDZ (n=21) boten keine Qualifizierungsmaßnahmen an, während elf FDZ zwischen einer und fünf Qualifizierungsmaßnahmen veranstalteten. Anzumerken ist, dass ein FDZ in der Lage war, mehr als 10 Qualifizierungsmaßnahmen, mehrheitlich in Form von Onlinekursen, anzubieten.

Die Anzahl der angebotenen Qualifizierungsmaßnahmen pro FDZ variiert stark

Abb. 23: Anzahl der Qualifizierungsmaßnahmen pro FDZ

Geclustert, n=38



© RatSWD 2022

Qualitätssicherung der Serviceangebote

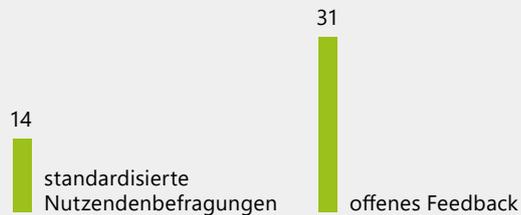
Fast alle FDZ führen prozessbegleitend Maßnahmen durch, um die Qualität ihrer Serviceangebote zu sichern (► **Abb. 24**). 31 FDZ ziehen ein offenes Feedback zur Qualitätssicherung heran, insgesamt 14 FDZ nutzen eine standardisierte Befragung der Datennutzenden (davon elf zusätzlich zum Feedback). Die Regelmäßigkeit solch einer Befragung unterscheidet sich zwischen den FDZ deutlich. Während sechs FDZ angeben, eine Befragung kontinuierlich durchzuführen, werden bei fünf FDZ Datennutzende seltener als einmal jährlich befragt.

Weiteres Feedback erhalten einige FDZ zudem über Feedback-Fragebögen im Anschluss an Schulungen sowie interne Evaluierungen seitens eines wissenschaftlichen Beirats.

Abb. 24: Qualitätssicherung durch Nutzendenbefragungen

Wie wird dieses Serviceangebot evaluiert bzw. die Qualität gesichert?
(Mehrfachantworten möglich)

■ Anzahl der FDZ in 2020 (n=38)



Qualitätssicherung findet mehrheitlich durch offenes Feedback der Nutzenden statt

Das erhaltene Feedback wird von den FDZ diskutiert, bewertet sowie wenn möglich umgesetzt und dient somit als wichtiger Anstoß für Weiterentwicklungen der Infrastruktur. Zudem wird vereinzelt Feedback in anonymisierter Form auf der Website veröffentlicht. Zusammenfassend zeigt sich, dass das auf unterschiedliche Weise erhaltene Feedback bei den FDZ eine bedeutsame Grundlage für Optimierungen des Serviceangebots wie auch für Verbesserungen interner Arbeitsprozesse ist.

6 Aktuelle Entwicklungen der Forschungsdateninfrastruktur in den Sozial-Verhaltens- und Wirtschaftswissenschaften

Die aktuellen Entwicklungen umfassen die internationale Ausrichtung der FDZ, sowohl im Bereich des Datenangebots als auch bei Kooperationen. Zudem werden Innovationen und Weiterentwicklungen im Bereich der Forschungsdateninfrastruktur vorgestellt.

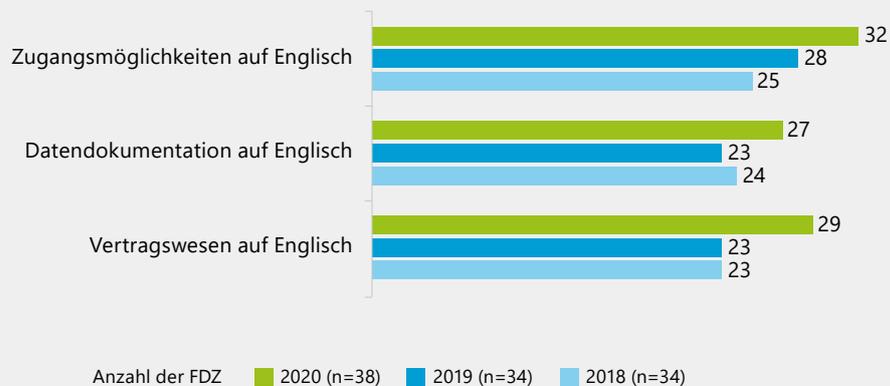
Internationalisierung

In der Wissensgesellschaft des 21. Jahrhunderts, deren globalen Verflechtung sich durch die digitale Transformation rasant beschleunigt, ergeben sich stetig neue Forschungsfelder, die nur im Rahmen einer internationalen Zusammenarbeit nachhaltig bearbeitet werden können. Die Voraussetzung dafür ist zum einen der Zugang internationaler Forschender zu den Daten in den jeweiligen Ländern. Zum anderen werden internationale Mikrodaten nachgefragt, die eine bessere vergleichende Analyse erlauben. Ein internationaler Austausch und Kooperationen gehören längst zum wissenschaftlichen Alltag und prägen auch die FDZ im zunehmenden Maße. So wurden zum Beispiel Zugangsmöglichkeiten und Dokumentationen für die internationale Wissenschaftscommunity in englischer Sprache geschaffen.

Abb. 25: Unterstützung internationaler Wissenschaftler:innen

In welcher Form unterstützen Sie Wissenschaftler:innen aus dem Ausland?
(Mehrfachantworten möglich)

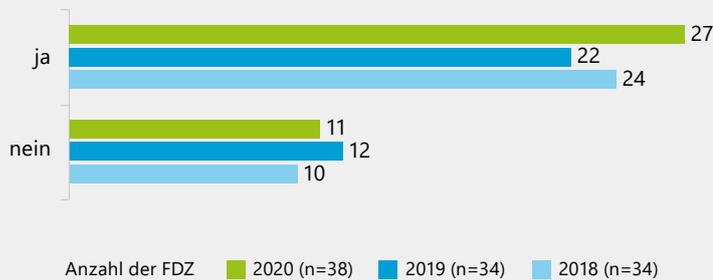
Vier weitere FDZ bieten Zugänge auf Englisch an



32 FDZ, und somit vier mehr als im Vorjahr, können einen Datenzugang auf Englisch anbieten. Dieser Anstieg ist insbesondere, aber nicht ausschließlich, durch die neu akkreditierten FDZ zu erklären. Im Vertragswesen wie auch in der Datendokumentation ist bei den FDZ ebenfalls eine Ausweitung des Angebots in englischer Sprache zu erkennen. Während im Vorjahr 23 FDZ das Vertragsmanagement sowie Datendokumentationen zusätzlich auf Englisch anboten, sind dies im Berichtsjahr 2020 29 bzw. 27 FDZ (► Abb. 25). Zusätzlich werden internationale Datennutzende durch eine persönliche Beratung telefonisch wie auch per E-Mail auf Englisch bzw. durch eine englischsprachige Version der Website und des Newsletters unterstützt. Weiterhin bieten einzelne FDZ regelmäßige Datenschulungen, Workshops und Konferenzen in englischer Sprache sowie Standorte im Ausland für Gastaufenthalte an.

Abb. 26: FDZ mit internationalen Kontakten

Bestehen Kontakte zu ausländischen Forschungseinrichtungen?



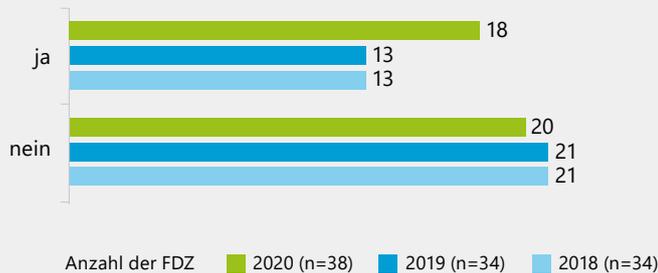
Kontakte zu ausländischen Forschungseinrichtungen nehmen zu

© RatSWD 2022

Die internationale Ausrichtung der FDZ zeigt sich zudem im bestehenden Austausch mit Forschungseinrichtungen im Ausland sowie in internationalen Forschungs Kooperationen nicht nur zur Beantwortung wissenschaftlicher Fragestellungen, sondern auch zur Abstimmung und Harmonisierung administrativer und datenschutzrechtlicher Regelungen. Insgesamt 27 FDZ pflegten im Berichtsjahr 2020 Kontakte zu internationalen Forschungseinrichtungen (► **Abb. 26**).

Abb. 27: FDZ mit engen internationalen Forschungs Kooperationen

Gibt es internationale Forschungs Kooperationen (d. h. wird gemeinsam an einem Thema gearbeitet)?



Internationale Forschungs Kooperationen nehmen zu

© RatSWD 2022

18 FDZ, und somit fast die Hälfte aller FDZ, sind im Rahmen einer Forschungs Kooperation im Austausch mit ausländischen Einrichtungen (► **Abb. 27**). Die Art der Forschungs Kooperation reicht von Projekten bis zu Arbeitsgruppen mit europäischen und internationalen Einrichtungen wie Universitäten, FDZ, Datenarchiven und Forschungsinstituten.

Innovationen und Neuerungen zur bestehenden Dateninfrastruktur

Die Innovationen und Neuerungen decken auch im Berichtsjahr 2020 thematisch wieder ein breites Spektrum ab. Hierzu gehört zunächst die stete Erweiterung des Datenangebots um neue Datensätze, die u. a. die Bereitstellung verknüpfter Datensätze von verschiedenen FDZ umfasst. Darüber hinaus wurden, wie auch schon im Vorjahr zu beobachten, in mehreren FDZ die bestehenden Datenzugangswege um die Möglichkeit einer kontrollierten Datenfernverarbeitung für einen erleichterten Datenzugang zu formal anonymisierten Mikrodaten erfolgreich erweitert.

Auch die Möglichkeiten für internationale Datennutzende mit Mikrodaten der FDZ wissenschaftliche Forschungsfragen zu bearbeiten, ist durch die Erschließung neuer Standorte im Ausland für Gastaufenthalte erweitert worden. Zudem trieben viele FDZ die Automatisierung von internen Arbeitsprozessen wie auch die Konzeption von Schulungen und Workshops als Online-Veranstaltungen voran.

Weiterentwicklung der Infrastruktur

Alle FDZ eint der Zustand der ständigen Weiterentwicklung. Das Datenangebot wird stetig vergrößert, die Möglichkeiten des Datenzugriffs werden ausgebaut, Serviceleistungen werden um Online-Angebote erweitert und die internationale Ausrichtung wird intensiviert. Getreu dem Motto „Stillstand ist Rückschritt“, beschäftigen sich alle FDZ auch weiterhin mit dem qualitativen wie quantitativen Ausbau des schon Erreichten. Hierbei ist insbesondere die heterogene Situation zwischen den FDZ zu beachten. Während einige FDZ bestimmte Serviceleistungen oder interne Prozesse schon über mehrere Jahre anbieten bzw. anwenden, befinden sich diese in anderen FDZ noch im Aufbau oder in der Planung. Daher überrascht es nicht, dass sich die Themen, zu denen seitens der FDZ Beratungsbedarf gesehen wird bzw. Interesse an einem Wissensaustausch vorhanden ist, mit den Themen vom Vorjahr größtenteils überschneiden.

Datenschutz ist auch im Berichtsjahr 2020 wieder ein Thema, zu dem viele FDZ Interesse an einem Wissensaustausch und Bedarf an Beratung zeigen. Hierunter fallen Themen wie verbindliche rechtliche Informationen zum Datenschutz und die Weiterentwicklung von Anonymisierungsmethoden vor dem Hintergrund der Anforderungen der EU-DSGVO.

Im Bereich *Forschungsdatenmanagement* lassen sich folgende Hauptthemen identifizieren, die für die FDZ von Interesse sind:

- Im Bereich des Datenzugangs wird ein Austausch zu den Themen alternativer und innovativer Datenzugangswege, insbesondere zu den Standards und den technischen Grundlagen für Remote Execution und Remote Desktop, gewünscht. Hierzu besteht zudem Austauschbedarf zur möglichen Automatisierung der Prüfung zur statistischen Geheimhaltung (Outputkontrolle).
- Im Bereich der Nutzendenverwaltung ist eine Definition des Begriffs wissenschaftliche Einrichtung mit einer entsprechenden Positiv- und Negativliste gewünscht. Zudem besteht Interesse an einem Austausch zu aktuellen Methoden und Werkzeugen des Anfrage- und Vertragsmanagements, insbesondere zu den Möglichkeiten einer Automatisierung administrativer Prozesse.
- Diskussionsinteresse besteht zudem zu Methoden und Werkzeugen zur (Meta-)Datenharmonisierung sowie zur Erleichterung des Informationsflusses in Form von strukturierten Metadaten von der Datenerhebung bis hin zur Endnutzung.
- Ebenfalls genannte Themen sind die Entwicklung von Kriterien für die Datenqualität, Maßnahmen zur Steigerung der Sekundärnutzungen, die Entwicklung von Portfolios sowie von Tools für die Unterstützung der Datenvorbereitung im Forschungsprozess und die Qualifikation von Nachwuchskräften.
- Auch der Umgang mit neuen Aufgaben, wie zum Beispiel die Rolle von FDZ als Datentreuhänder, wurde thematisiert.

7

Fokusthema: Metadatenstandards

Die sogenannten **FAIR-Prinzipien (kurz für: Findability, Accessibility, Interoperability, Re-usability)** haben sich in den letzten Jahren als Leitlinien für die Gestaltung von Forschungsdateninfrastrukturen etabliert. Sie werden national¹⁷ und international¹⁸ als Standards für das Teilen von Daten angesehen. Im Vordergrund der Diskussion steht dabei häufig die Möglichkeit, Daten nachzunutzen. Eine Bedingung dafür, dass Daten gewinnbringend nachgenutzt werden können, sind hochwertige und granulare Metadaten. Daten, die unvollständig oder unverständlich beschrieben sind, können auch bei einem freien Zugang nicht verwendet werden.

Mit dem Start des KonsortSWD innerhalb der NFDI wurde eine Struktur geschaffen, um die FAIRness von sozialwissenschaftlichen Daten noch weiter zu erhöhen, indem die Metadatenqualität verbessert und die Erzeugung von Metadaten noch weiter aufeinander abgestimmt wird. Eine der drei Säulen des KonsortSWD ist die Stärkung der FAIRness von Daten und Metadaten.

Da das Netzwerk aus FDZ in den vergangenen Jahren stark gewachsen ist, wurde es notwendig, eine empirische Grundlage zu generieren, wie Metadaten in den FDZ erzeugt und verwaltet werden. Daher hat der FDI Ausschuss beschlossen, das Schwerpunktmodul im Monitoring des Jahres 2020 den Metadatenstandards zu widmen.

Die Darstellung der Metadatenpraktiken in den FDZ erfolgt entlang der vier einzelnen Säulen der FAIR-Prinzipien und schließt jeweils mit einer kurzen Einschätzung zum Stand der Umsetzung in den FDZ ab.



¹⁷ Hartl, N., Wössner, E., & Sure-Vetter, Y. (2021). Nationale Forschungsdateninfrastruktur (NFDI). *Informatik Spektrum*, 44(5), 370-373. <https://doi.org/10.1007/s00287-021-01392-6>.

¹⁸ Streit, A., & Wezel, J. (2021). Deutschland in der European Open Science Cloud. In M. Putnigs, H. Neuroth & J. Neumann (Hg.), *Praxishandbuch Forschungsdatenmanagement* (S. 31-52). De Gruyter Saur. <https://doi.org/10.1515/9783110657807-003>.

Findability (Auffindbarkeit)

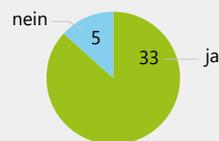
Kernelement FAIRer Dateninfrastrukturen ist die systematische Verwendung von PID für Daten und Metadaten. Diese Identifikatoren erlauben es, den Datensatz und dessen Metadaten auffindbar und zitierbar zu machen. Im Gegensatz zu anderen seriellen Identifikatoren (bspw. URL-Adressen) löst ein PID auf das Objekt selbst auf und nicht auf seinen Standort im Internet. Ändert sich der Standort eines mit einem PID referenzierten digitalen Objekts, so bleibt der Identifikator derselbe. Werden PID verwendet, ist damit in der Regel ein Minimum an generischen Metadaten zum referenzierten Objekt verbunden (zum Beispiel Autorenschaft, Titel, Erhebungszeiträume etc.).

In den vergangenen Jahren war die Einführung von PID einer der Schwerpunkte in der Weiterentwicklung der FDZ. Von den 38 akkreditierten verwenden 33 bereits PID, davon 31 die DOI von DataCite¹⁹ und zwei auch sogenannte *Handles*²⁰. Ein FDZ verwendet selbst entwickelte PID (► **Abb. 28**). Drei weitere FDZ arbeiten gerade an der Einführung von DOI für die Daten.

Abb. 28: Verwendung von persistenten Identifikatoren

Welche weltweit eindeutigen, dauerhaften und auflösbaren Identifikatoren (im Folgenden als Persistenter Identifier – PID – bezeichnet) verwendet Ihr FDZ, um Daten und Metadaten zu referenzieren? n=38

Das FDZ verwendet mindestens einen Identifikator um Daten und Metadaten zu referenzieren.



davon (n=33, Mehrfachantworten möglich*)

Digital Object Identifier (DOI)	31
Handle (hdl)	2
Eigenentwicklung	1

* Weitere Antwortmöglichkeiten, die von keinem FDZ angegeben wurden: Persistent Uniform Resource Locator (PURL), Uniform Resource Name (URN)

© RatSWD 2022

PID werden in den FDZ für die Referenzierung von Daten verwendet, wobei die konkrete Umsetzung sich unterscheidet. In 20 Fällen werden PID für alle Daten aus einer Studie vergeben (d. h. ein PID löst ggf. mehrere Datenfiles auf, die zu einer Studie gehören, wie dies bei einzelnen Panelwellen der Fall ist). In 19 FDZ werden die PID für die einzelnen Datenfiles vergeben (jede Panelwelle bekommt eine eigene PID) und in fünf FDZ werden PID auch für einzelne Datenobjekte vergeben (zum Beispiel Interviewtranskripte). Für einzelne Objekte oder Fragmente innerhalb von Datenfiles werden in den FDZ keine PID vergeben, wohl aber in mindestens acht FDZ für Kontext- oder Studiendokumente (inkl. Fragebögen, Codebücher, Methodenberichte, Working Paper und Testinstrumente). In 30 FDZ lösen die PID nicht direkt auf das digitale Objekt auf, sondern auf eine sogenannte *Landing Page*, die Metadaten über die Daten und Dokumente enthält und auf der Hinweise für den Datenzugriff zu finden sind. Nur in einem FDZ lösen die PID direkt auf die digitalen Objekte auf.

Wie in den Best Practices zur Verwendung von PID vorgesehen, werden diese von 29 FDZ in den Metadaten explizit aufgeführt (Zitierlink der Daten). So ist die Beziehung zwischen PID und Daten für Nutzende zu erkennen. Die Verknüpfung der PID mit den Metadaten erfolgt technisch über die Metadaten schemata von DOI in *dajra*²¹ oder *DublinCore*²². In einem Fall wird *CMDI*²³ verwendet.

¹⁹ <https://datacite.org/> (Zugriff am 31.01.2022).

²⁰ <https://www.handle.net/index.html> (Zugriff am 31.01.2022).

²¹ *dajra* bezeichnet die Registrierungsagentur für sozial- und wirtschaftswissenschaftliche Forschungsdaten des Leibniz-Institut für Sozialwissenschaften (GESIS) und des Leibniz-Informationszentrum Wirtschaft (ZBW). <https://www.da-ra.de/> (Zugriff am 31.01.2022).

²² *Dublin Core* ist ein Metadatenstandard zur Beschreibung von elektronischen Ressourcen, der von der *Dublin Core Metadata Initiative (DCMI)* spezifiziert wurde. <https://dublincore.org/> (Zugriff am 31.01.2022).

²³ *Component Metadata Infrastructure (CMDI)* ist eine komponentenbasierte Metadatenbeschreibung. <https://www.clarin.eu/content/component-metadata> (Zugriff am 31.01.2022).

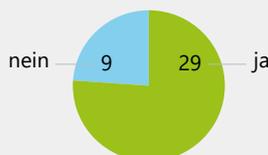
Während sich PID für ganze Datensätze bzw. Studien in den FDZ bewährt haben, sieht das Konzept der *FAIR Digital Objects* (ein Modell für eine technische Umsetzung der FAIR-Prinzipien im Datenmanagement) die Verwendung von PID auch für Variablen, Fragmente oder andere Attribute in Datendateien vor.²⁴ Eine feingranulare Referenzierung von Daten und Metadaten über PID soll in Zukunft den automatisierten Zugriff auf (Meta-)Daten erleichtern. Tatsächlich berichten sechs FDZ, die Verwendung weiterer PID für eine granulare Referenzierung von Datenobjekten zu planen. Anders als bei den PID für Daten bzw. Studien sollen für die Attribute allerdings keine DOI, sondern Uniform Resource Names (URN) verwendet werden. Der Grund ist vermutlich, dass die Vergabe einer sehr großen Zahl von DOI zu hohe Kosten verursachen würde, solange keine „Nationallizenz“ für DOI im Forschungsdatenmanagement existiert.

Eine weitere Säule für die verbesserte Auffindbarkeit von Daten ist die Beschreibung der Daten mit ausführlichen Metadaten. Damit solche Beschreibungen maschinenlesbar und untereinander verknüpfbar sind, sollen Metadaten anhand eines etablierten Standards generiert werden. ► **Abb. 29** zeigt, dass die FDZ in der großen Mehrzahl bereits Metadatenstandards für die Dokumentation einsetzen und zum Teil auch verschiedene Standards kombinieren, insbesondere wenn ein eigenes Metadatenschema für die Datenbeschreibung entwickelt wurde. 20 FDZ setzen das Metadatenschema von *da|ra* ein, das für die Vergabe von DOI benötigt wird. Jeweils elf FDZ verwenden für die Dokumentation ein eigenes Metadatenschema oder den Metadatenstandard der Data Documentation Initiative (DDI). Das DDI-Metadatenschema wird von den FDZ an ihre Bedürfnisse angepasst. Ein Beispiel dafür ist das bei GESIS verwendete DDI-LimDAS²⁵ oder das Metadatenmodell des FDZ-BO²⁶. Weitere Metadatenstandards, die von den FDZ verwendet werden, sind DublinCore, Schema.org, PsychData, Inexda, die CLARIN-Component MetaData Infrastructure, das Metadatenstandard des Verbundes Forschungsdaten Bildung und INSPIRE-OGC.

Abb. 29: Verwendung von Metadatenstandards und -schemata

Welchen Metadatenstandard oder welches Metadatenchema verwendet Ihr FDZ, um die Auffindbarkeit der Daten sicherzustellen? n=38

Das FDZ verwendet mindestens einen Standard oder ein Metadatenchema, um die Auffindbarkeit der Daten sicherzustellen.



davon (n=29, Mehrfachantworten möglich*)

DataCite oder <i>da ra</i> Metadata Schema	20
Data Documentation Initiative (DDI)	11
Ein eigenes Metadatenchema	11
Dublin Core (DC)	8
Schema.org	5

* Weitere Antwortmöglichkeiten, die von keinem FDZ angegeben wurden: Statistical Data and Metadata Exchange (SDMX), Directory Interchange Format (DIF), Qualitative Data Exchange Format (QuDEx) oder REFI-QDA, Data Catalog Vocabulary (DCAT)

Die Auffindbarkeit von Daten wird durch die Weitergabe von Metadaten an spezialisierte Suchportale erhöht. Metadaten der 28 FDZ, die für die Dokumentation der Daten die Angebote von *da|ra* nutzen, werden von internationalen Datensuchportalen wie zum Beispiel der DataCiteSearch, dem CESSDA Data Catalogue, dem europäischen Suchportal B2Find, dem Suchportal OpenAire und der GoogleDatasetSearch indiziert. Mindestens vier von diesen FDZ bereiten die Metadaten speziell auch für das Suchportal des Verbundes Forschungsdaten Bildung auf. Drei FDZ, die nicht *da|ra* nutzen, sind in der DataCiteSearch vertreten, und drei FDZ in anderen Suchportalen. Nur drei FDZ speisen ihre Metadaten (noch) nicht in internationale Suchportale für Forschungsdaten (wie unter anderem B2Find, OpenAire, CLARIN-VLO).

24 Betancort Cabrera, N., Bongartz, E. C., Dörrenbächer, N., Goebel, J., Kaluza, H., & Siegers, P. (2020). *White Paper on implementing the FAIR principles for data in the Social, Behavioural, and Economic Sciences*. RatSWD Working Paper 274/2020. Rat für Sozial- und Wirtschaftsdaten (RatSWD). <https://doi.org/10.17620/02671.60>.

25 <http://dx.doi.org/10.21241/ssoar.65593> (Zugriff am 31.01.2022).

26 https://www.diw.de/documents/publikationen/73/diw_01.c.620524.de/diw_datadoc_2019-099.pdf (Zugriff am 31.01.2022).

Zwischenfazit Findability

Persistente Identifikatoren und Metadaten nach etablierten Standards sind die Eckpfeiler für eine FAIRe Dateninfrastruktur. Die akkreditierten FDZ sind hier als gute Vorbilder vorangegangen. Für die Nutzenden sind die Daten der FDZ bis auf wenige Ausnahmen über fachspezifische und generische Suchportale auffindbar. Neun FDZ verwenden kein Metadatenschema und zehn FDZ verwenden ein selbst entwickeltes Metadatenschema für die Beschreibung ihrer Daten. Die Heterogenität in den Datenbeschreibungen spiegelt die Vielfalt der Domänen wider, für die die FDZ Daten anbieten. Die FDZ, die noch keine PID verwenden, sollten deren Einführung planen, um für neue Entwicklungen in der NFDI oder der European Open Science Cloud (EOSC) anschlussfähig zu sein und die eigenen Daten sichtbarer zu machen. Die wenigen FDZ, die ihre Metadaten noch nicht an Suchportale weitergeben, sollten eine technische Lösung hierfür anstreben, damit Nutzende die Daten besser finden können. In Zukunft ist zu prüfen, wie eine stärkere Standardisierung der Metadatenpraxis in den FDZ erfolgen kann, um die automatisierte Verarbeitung von Metadaten zu optimieren.

Accessibility (Zugänglichkeit)

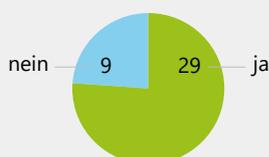
Zugänglichkeit im Sinne der FAIR-Prinzipien meint, dass Metadaten über standardisierte Kommunikationsprotokolle erreichbar und für Menschen und Maschinen abrufbar sind. Daten und Metadaten sollen langfristig erhalten bleiben. Die verwendeten Protokolle sollen offen, frei und universell implementierbar sein. Das bedeutet, dass für die Nutzung der Kommunikationsprotokolle keine Zusatzkosten entstehen sollen.

Da die Außendarstellung der FDZ über die Webseiten der Institutionen erfolgt, sind eigene Webseiten basierend auf dem html- oder https-Protokoll die am weitesten verbreiteten Kommunikationsprotokolle der FDZ. Darüber hinaus verwenden acht FDZ das Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), das speziell für den Austausch von Metadaten von Archiven entwickelt wurde und sieben FDZ nutzen eine Representational State Transfer-Application Programming Interface (REST-API) – eine generische Schnittstelle für den Austausch von Daten in Webanwendungen (► **Abb. 30**). OAI-PMH und REST-API sind für einen automatisierten Datenaustausch konzipiert und sorgen dafür, dass die Metadaten auch von Maschinen verarbeitet werden können. Die OAI-PMH wird auch von *da|ra* angeboten, sodass die von den FDZ an *da|ra* gemeldeten Metadaten über die *da|ra* API abgerufen werden können.

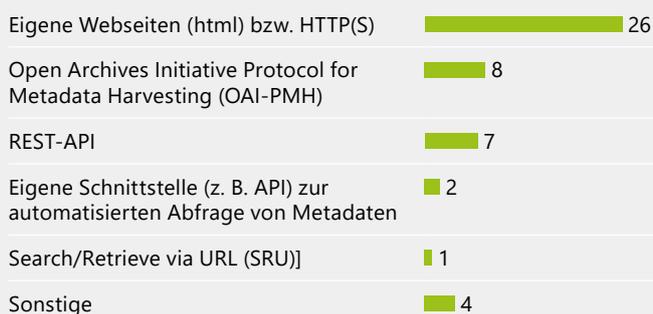
Abb. 30 Standardisierte Kommunikationsprotokolle für den Metadatensatz

Welches standardisierte Kommunikationsprotokoll verwendet Ihr FDZ für den Metadatensatz?
n=38 (Mehrfachantworten möglich)

Das FDZ verwendet mindestens ein standardisiertes Ergebnisprotokoll



davon (n=29, Mehrfachantworten möglich*)



*Weitere Antwortmöglichkeiten, die von keinem FDZ angegeben wurden: File Transfer Protocol (FTP) und Simple Mail Transfer Protocol (SMTP)

Die Frage der Authentifizierung von Nutzenden und die Verwaltung von Nutzungsrechten sind für die FDZ aufgrund der verschiedenen Schutzklassen von Daten und Metadaten eine entscheidende Aufgabe. Die FAIR-Prinzipien fordern einen möglichst offenen Zugang zu den Daten, soweit dies rechtlich möglich ist. Die überwiegende Mehrzahl der FDZ gewährt einen freien Zugang zu Metadaten und Studiendokumentationen. In sechs FDZ müssen sich Nutzende authentifizieren, bevor Metadaten verfügbar sind (► **Abb. 31**).

Abb. 31: Offener Zugang zu Metadaten

Verwenden Sie für Ihr FDZ eine Authentifizierungsinfrastruktur (Benutzendenkonto), um den Zugriff auf Metadaten und Studiendokumentationen zu regulieren? n=38

Nein, der Zugriff auf alle Metadaten und Studiendokumentationen ist nicht reguliert		29
Ja, für einen Teil der Metadaten und Studiendokumentationen		4
Ja, für alle Metadaten und Studiendokumentationen		2
keine Angabe		3

© RatSWD 2022

Eine weitere Dimension der Zugänglichkeit ist die langfristige Sicherung von Metadaten, damit das Wissen über verfügbare Datenbestände nicht verloren geht. Nur sieben FDZ haben keine Strategie für den langfristigen Erhalt der Metadaten über die Existenz des FDZ hinaus. Die anderen FDZ übergeben die Metadaten an Infrastrukturpartner, die deren langfristige Verfügbarkeit sicherstellen können. Genannt werden da|ra und das GESIS Datenarchiv, das selbst allerdings noch nicht über einen offiziellen Nachfolgeplan verfügt. Andere Metadatendienste sind HEBIS²⁷, Pangea²⁸, PsychArchives²⁹ oder Dienste von Universitäten.

Zwischenfazit Accessibility

Die Zugänglichkeit zu Metadaten in den FDZ entspricht weitgehend den Ansprüchen der FAIR-Prinzipien. In allen Fällen liegen Informationen zum Datenbestand und Datenzugang vor. 15 FDZ stellen die Metadaten über standardisierte Programmierschnittstellen auch maschinenlesbar bereit. Damit können Suchportale die Metadaten in die eigenen Suchindices übernehmen und tragen so auch zur langfristigen Sicherung der Metadaten bei. Durch eine Ausweitung der Verwendung von Programmierschnittstellen für den Zugriff auf Metadaten könnte die Zugänglichkeit in Zukunft noch verbessert werden. Parallel dazu könnten – z. B. im Zusammenhang mit CoreTrustSeal-Zertifizierungen – reziproke Nachfolgevereinbarungen für (Meta-)Daten zwischen den FDZ getroffen werden, um das Ausfallrisiko weiter zu reduzieren.

27 <https://www.hebis.de/> (Zugriff am 31.01.2022).

28 <https://www.pangaea.de/> (Zugriff am 31.01.2022).

29 <https://www.psycharchives.org/> (Zugriff am 31.01.2022).

Interoperability (Interoperabilität)

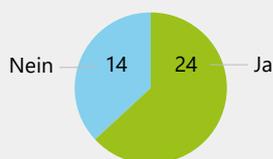
Interoperabilität ist letztlich das anspruchsvollste der FAIR-Prinzipien, weil es darauf abzielt, Daten so zu erschließen, dass Schnittstellen für deren Verknüpfung entstehen. Verknüpft werden soll jedoch nicht nur durch menschliche Nutzende, sondern auch durch Computersysteme. Dafür ist die Verwendung weit verbreiteter Standards für die inhaltliche und die granulare Beschreibung von Daten zentral. Ziel ist es, dass Computersysteme entscheiden können, ob der Inhalt der Datensätze vergleichbar ist. Für die Erstellung und Anwendung solcher Metadaten werden kontrollierte Vokabulare, Ontologien oder Thesauri und ein klar definiertes Framework benötigt, z. B. im Sinne des Semantic Web.

Solche Semantic Web-Technologien werden von den FDZ in Ansätzen bereits verwendet (► **Abb. 32**). 17 FDZ stellen ihre Metadaten als Extensible Markup Language (XML) Format bereit. Der Vorteil dieser Darstellung ist, dass sie sowohl für Menschen als auch Computersysteme interpretierbar ist. Drei FDZ verwenden die JavaScript Object Notation (JSON LD). Dieser Standard wird für den Austausch von Daten zwischen verschiedenen Anwendungen verwendet. Ein FDZ verwendet das Resource Description Framework (RDF), das vor allem im Bereich der Webentwicklung weite Anwendung findet. 14 FDZ stellen ihre Metadaten in keiner Sprache der Wissensrepräsentation dar. Dies schränkt Möglichkeiten einer automatisierten Evaluation und Verknüpfung von Daten ein.

Abb. 32: Verwendete Sprache/n der Wissensrepräsentation zur Beschreibung und Darstellung von Metadaten

Welche Sprache/n der Wissensrepräsentation, die Maschinenlesbarkeit erlauben, verwenden Sie in Ihrem FDZ zur Beschreibung und Darstellung von Metadaten? n=38

Das FDZ verwendet zur Beschreibung und Darstellung von Metadaten mindestens eine Sprache der Wissensrepräsentation, die Maschinenlesbarkeit erlaubt.



davon (n=24, Mehrfachantworten möglich*)

Extensible Markup Language (XML)	17
JavaScript Object Notation basierte Datenverlinkung (JSON LD)	3
Ressource Description Framework (RDF)	1
Sonstige	2
keine Angabe	4

*Weitere Antwortmöglichkeiten, die von keinem FDZ angegeben wurden: Resource Description Framework Schema (RDFS), RDF(S) extensible knowledge representation model, Web Ontology Language (OWL), Agent Markup Language-Programm + Ontology Inference Layer (DAML+OIL)

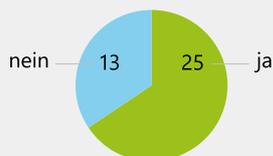
Für eine standardisierte inhaltliche Erschließung ist nicht nur die Verwendung von offenen Austauschformaten notwendig, sondern auch eine semantisch äquivalente Beschreibung der Inhalte. Dafür müssen Messungen gleicher oder ähnlicher Konstrukte in der Datenbeschreibung mit den gleichen Begriffen (Schlagwörtern) erfasst werden. Zu diesem Zweck werden in Bibliotheken und Archiven traditionell Thesauri und Ontologien verwendet, die eine einheitliche Verschlagwortung sicherstellen sollen. Über die Vergabe ähnlicher Schlagwörter können verknüpfbare Inhalte identifiziert werden. Auch wenn verschiedene Thesauri eingesetzt werden, helfen Mappings dabei, vergleichbare Inhalte in den Daten zu identifizieren.

► **Abb. 33** zeigt die Vielfalt der verwendeten Thesauren auf. Am weitesten verbreitet sind die kontrollierten Vokabulare der Data Documentation Initiative (DDI CV), die vor allem für Standardfelder in der Dokumentation von Befragungsstudien entwickelt wurden (z. B. Erhebungsmethode, Stichprobenverfahren). Auch der Thesaurus Sozialwissenschaften (TheSoz)³⁰, die CESSDA Topic Classification³¹ und die Klassifikation des Journal of Economic Literature (JEL)³² werden von einigen FDZ verwendet. Daneben dominieren fachspezifische Thesauren, wie OGC-INSPIRE³³ oder PSYNDEX Terms³⁴ oder das kontrollierte Vokabular des Verbundes Forschungsdaten Bildung³⁵. Die FDZ bei GESIS verwenden zudem die ISO Normvokabulare 639-3 für Sprachen und 3166-1/2/3 für die Verschlagwortung von Ländern und Regionen. Allerdings verwenden 13 FDZ keine kontrollierten Vokabulare für die inhaltliche Erschließung ihrer Daten und neun FDZ haben eigene Vokabulare entwickelt.

Abb. 33: Beschreibung von Datensätzen durch Vokabulare, Ontologien oder Thesauri

Mit welchen kontrollierten Vokabularen, Ontologien oder Thesauri, die zur Beschreibung von Datensätzen verwendet werden, annotieren Sie in Ihrem FDZ die Metadatendokumentation? n=38

Das FDZ verwendet zur Beschreibung von Datensätzen mindestens ein kontrolliertes Vokabular, eine Ontologie oder einen Thesaurus.



davon (n=25, Mehrfachantworten möglich*)

Selbst entwickelte kontrollierte Vokabulare, Ontologien oder Thesauri	9
Kontrollierte Vokabulare der Data Documentation Initiative (DDI CV)	9
Thesaurus Sozialwissenschaften (TheSoz)	6
Klassifikation des Journal of Economic Literature (JEL)	3
European Language Social Science Thesaurus (ELSST)	1

*Weitere Antwortmöglichkeit, die von keinem FDZ angegeben wurde: Standard Thesaurus Wirtschaft (STW)

Für die Erzeugung standardisierter Metadaten werden geeignete Werkzeuge benötigt, die relevante Metadatenstandards unterstützen, die Vergabe standardisierter Schlagworte aus kontrollierten Vokabularen sowie den Zugriff über die oben genannten Semantic Web Technologien ermöglichen. Hier zeigt die Befragung der FDZ, dass bislang nur in Ausnahmefällen auf etablierte Editoren zurückgegriffen wird. Die meisten FDZ verwenden eigene Metadateneditoren (Studienebene: n=16; Variablenebene: n=9) oder keine Editoren (Studienebene: n=13; Variablenebene: n=20). Kommerzielle oder OpenSource Editoren für Metadaten werden fast gar nicht verwendet.³⁶

Positiv hervorzuheben ist allerdings, dass die Dokumentation der Daten in den meisten Fällen Verweise auf die mit den Daten verbundenen Publikationen enthält (► **Abb. 34, S. 44**). Das betrifft entweder Publikationen, die weitergehende Informationen zu Erstellung der Daten berichten, oder solche, die aus der Auswertung der Daten hervorgegangen sind.

30 <https://lod.gesis.org/thesoz/de/> (Zugriff am 31.01.2022).

31 <https://vocabularies.cessda.eu/vocabulary/TopicClassification> (Zugriff am 31.01.2022).

32 <https://www.aeaweb.org/jel/guide/jel.php> (Zugriff am 31.01.2022).

33 <https://inspire.ec.europa.eu/> (Zugriff am 31.01.2022).

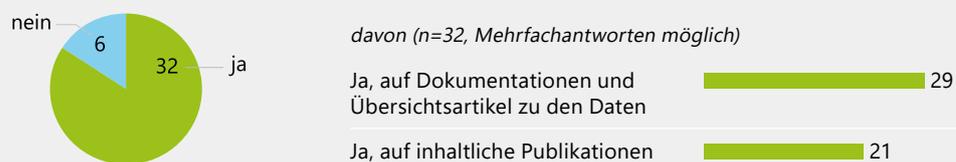
34 <https://www.psyndex.de/ueber/inhalte-aufbau/schlagwoerter-klassifikationen/> (Zugriff am 31.01.2022).

35 https://www.forschungsdaten-bildung.de/files/fdbinfo_8_metadatenset_v1.0.pdf (Zugriff am 31.01.2022).

36 Zu den wenigen genannten Ausnahmen gehören Colectica, Questasy, GeoMIS, DSpace-Web-UI und da|ra.

Abb. 34: Referenzierung auf Publikationen in den Metadaten

Verweisen Sie in Ihrem FDZ in den Metadaten auf verbundene Publikationen? n=38



© RatSWD 2022

Zwischenfazit Interoperabilität

Um eine hohe Interoperabilität von Forschungsdaten zu erreichen, müssen die Metadaten stark standardisiert sein und in maschinenlesbaren Austauschformaten vorliegen. Sie müssen eine semantisch äquivalente inhaltliche Beschreibung der Daten vornehmen. Das sind hohe Ansprüche, die harmonisierte Metadatenpraktiken voraussetzen. Zwar beruht die Datendokumentation der FDZ in weiten Teilen bereits auf gängigen Standards (siehe oben), aber das Angebot von Metadaten in modernen Austauschformaten stellt weiterhin die Ausnahme dar. Um die Verknüpfung der (Meta-)Daten aus den FDZ zu vereinfachen, könnte eine Harmonisierung der Metadatenpraxis über gemeinsame Metadateneditoren eine Lösung darstellen. Da einsatzfähige OpenSource Systeme derzeit nicht verfügbar sind, ist ein solches Ziel kurz- und mittelfristig nicht zu erreichen. Empfohlen werden kann allerdings die systematischere Verwendung von kontrollierten Vokabularen innerhalb der in den FDZ implementierten Standards, um die semantische Erschließung der Daten zu vereinheitlichen.

Re-usability (Wiederverwendbarkeit)

Die Nachnutzung von Daten setzt voraus, dass die Nutzenden nachvollziehen können, wie die Daten entstanden sind und welche Auswertungen mit den Daten zulässig sind. Das Prinzip der Wiederverwendbarkeit unterscheidet sich von der Auffindbarkeit vor allem darin, dass hier die Fähigkeit eines Nutzenden (Maschine oder Mensch) gefragt ist, zu entscheiden, ob spezifische Daten in ihrem jeweiligen Kontext dem eigenen Nutzungszweck gerecht werden. Um diese Entscheidung zu treffen, sind Metadaten erforderlich, die den Kontext der Datengenerierung beschreiben und transparent machen, wer unter welchen Bedingungen die Daten nutzen darf. Dies gilt auch für die Nutzung der Metadaten, auf die sich die folgenden Ausführungen beziehen.

Nutzungsbedingungen für Metadaten können zum Beispiel über Standard-Lizenzmodelle festgelegt werden (für eine Übersicht der Standard-Lizenzmodelle siehe ▶ **Tab. 1**). ▶ **Abb. 35, S. 46** zeigt, dass 23 FDZ ihre Metadaten unter einem bestimmten Lizenzmodell zur Nachnutzung anbieten. Davon verwenden acht FDZ eigene Lizenzen und 15 verwenden verschiedene Varianten der Creative Commons Lizenzen. Am weitesten verbreitet ist eine vollkommen freie Nutzung (CC0), die auch kommerzielle Verwendungen miteinschließt. So können Metadaten auch von kommerziellen Suchportalen geharvestet und indiziert werden. Einige FDZ verwenden auch eine CC BY Lizenz, die eine Namensnennung bei der Verwendung verlangt. Nur fünf FDZ verwenden restriktivere Lizenzen, u. a. um eine kommerzielle Nutzung von Metadaten auszuschließen.

Tab. 1: Creative Commons Lizenzen³⁷

Lizenz	Erlaubt sind:	Unter der Bedingung:
CC BY	Vervielfältigung, Weitergabe, Erstellung von Bearbeitungen sowie deren Vervielfältigung und Weitergabe für kommerzielle und nicht-kommerzielle Zwecke	Namensnennung: Bezeichnung des Erstellers (soweit angegeben); Nennung des jeweiligen Lizenztyps und Referenz auf Lizenztext durch URI/Hyperlink; URI/Hyperlink zum lizenzierten Material, soweit vernünftigerweise praktikabel; Copyright-Vermerk und Hinweis auf Haftungsausschluss (beides nur, soweit angegeben); ggf. Hinweis, wenn lizenziertes Material verändert wurde
CC BY-SA	s.o.	Namensnennung (s. o.); Share Alike: abgewandelte Material muss unter vergleichbarer freier Lizenz zur Verfügung gestellt werden
CC BY-ND	Vervielfältigung, Weitergabe und Bearbeitung für kommerzielle und nicht-kommerzielle Zwecke, aber keine Vervielfältigung/Weitergabe von Bearbeitungen	Namensnennung (s. o.)
CC BY-NC	Vervielfältigung, Weitergabe, Erstellung von Bearbeitungen sowie deren Vervielfältigung und Weitergabe, aber nur für nicht-kommerzielle Zweck	Namensnennung (s. o.)
CC BY-NC-ND	Vervielfältigung, Weitergabe, Erstellung von Bearbeitungen sowie deren Vervielfältigung und Weitergabe, aber nur für nicht-kommerzielle Zwecke und keine Vervielfältigung/Weitergabe von Bearbeitungen	Namensnennung (s. o.)
CC0	z. T. Verzicht auf das Urheberrecht; da dies im deutschen Urheberrecht nicht möglich ist, weitest-mögliche Einräumung von Nutzungsrechten	grundsätzlich keine Namensnennung erforderlich
CC0 Plus (inoffizieller Lizenztyp, z. T. in der Bibliotheks- praxis genutzt*)	wie CC0	wie CC0, aber unverbindliche Aufforderung zur Namensnennung, soweit praktikabel

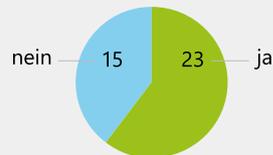
*Siehe z. B. den Usage Guide der British Library in Bezug auf Catalogue datasets, <https://www.bl.uk/about-us/terms-and-conditions/catalogue-datasets-in-rdf-and-csv>.

³⁷ Die Tabelle 1 ist ein unveränderter Nachdruck aus: Lauber-Rönsberg, A. (2021). Rechtliche Aspekte des Forschungsdatenmanagements. In M. Putnigs, H. Neuroth & J. Neumann (Hg.), *Praxishandbuch Forschungsdatenmanagement*. S. 89-114. De Gruyter Saur. <https://doi.org/10.1515/9783110657807>. Diese Veröffentlichung ist unter der Creative-Commons-Lizenz (CC BY 4.0) lizenziert: <https://creativecommons.org/licenses/by/4.0/> (Zugriff am 31.01.2022).

Abb. 35: Nutzungslizenzen der Metadatendokumentation

Welche Nutzungslizenz verwendet Ihr FDZ für die Metadatendokumentation? n=38

Das FDZ hat mindestens eine Nutzungslizenz festgelegt.



davon (n=23, Mehrfachantworten möglich*)

Eigene Lizenz	8
CC 0	6
CC BY	5
CC BY SA	2
CC BY NC	2
CC BY NC SA	1

* Weitere Antwortmöglichkeiten, die von keinem FDZ angegeben wurden: CC BY ND und CC BY NC ND

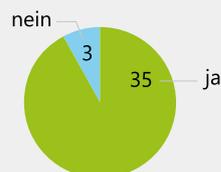
© RatSWD 2022

Um die Brauchbarkeit von Daten zu beurteilen, benötigen Nutzende detaillierte Informationen zu den datengenerierenden Verfahren und – im besten Fall – auch zu Prozessen der Datenverarbeitung, vor allem wenn Daten dabei verändert werden. In der Terminologie der FAIR-Prinzipien wird in diesem Zusammenhang von Provenienz-Informationen gesprochen.

Abb. 36: Provenienz-Informationen

Welche Provenienz-Informationen veröffentlicht Ihr FDZ im Rahmen der Studiendokumentation? n=38

Das FDZ veröffentlicht mindestens eine Provenienz-Informationen im Rahmen der Studiendokumentation.



davon (n=35, Mehrfachantworten möglich)

Methodenberichte	26
Fragebögen	23
Datenquelleninformationen	16
Anonymisierungskonzept	14
Datenqualitätsberichte	10
Aufbereitungsskripte	6
Transkriptionsmethoden	4
Rechteinformationen (z. B. beim Zusammenspielen von Daten oder beim Webscrapen)	2

© RatSWD 2022

► **Abb. 36** veranschaulicht die große Bandbreite an Provenienz-Informationen, die von den FDZ bereitgestellt werden. Welche Informationen für die Nachnutzung relevant sind, hängt stark von den Datentypen ab. So publiziert die Mehrheit der FDZ Methodenberichte oder Transkriptionsmethoden und Fragebögen. Letztere sind natürlich nur für Befragungsstudien relevant. Ebenfalls zu den häufig bereitgestellten Arbeitshilfen zählen Datenquelleninformationen und Anonymisierungskonzepte. Verfahrensbeschreibungen der Datenaufbereitung werden zum Teil als Technical Papers publiziert. Auch Datenqualitätsberichte gehören zu den Leistungen einiger FDZ.

Zwischenfazit Re-usability

Die Hürden bei der Verwendung von Metadaten der FDZ für wissenschaftliche Zwecke sind gering. In den meisten Fällen ist eine freie Verwendung möglich oder es müssen Quellen genannt werden. Dies entspricht den Empfehlungen für die Umsetzung der FAIR-Prinzipien in den Sozial- und Wirtschaftswissenschaften.³⁸ In Zukunft sollte geprüft werden, ob verstärkt CC-Lizenzen für die Nachnutzung von Metadaten aus den FDZ verwendet werden können, um die Transparenz aus Nutzendensicht zu erhöhen und die Nutzungsbedingungen auch für Maschinen verständlich zu machen.

Die Nachvollziehbarkeit von datengenerierenden Prozessen wird von den FDZ durch eine Vielzahl unterschiedlicher Dokumente sichergestellt. Gemeinsam ist allen, dass die Information noch nicht standardisiert in die Metadaten schemata integriert ist, sondern in der Regel aus den Dokumenten extrahiert werden muss. Bislang veröffentlichen nur 14 FDZ ihre Konzepte zur Anonymisierung von Daten. Da gerade in diesem Schritt der Datenaufbereitung die Daten regelmäßig stark verändert werden, kann ein transparentes Anonymisierungskonzept für die Nutzenden nachvollziehbar machen, welche Änderungen vorgenommen wurden. Wenn keine rechtlichen Gründe gegen eine Veröffentlichung der Anonymisierungskonzepte sprechen, sollten solche Informationen von den FDZ publik gemacht werden.

FAZIT: FAIRe (Meta-)Daten?

Qualitativ hochwertige Metadaten sind der Schlüssel zu einer FAIRen Infrastruktur für Forschungsdaten. Die FDZ verfügen mit ihrer langjährigen Erfahrung im Datenmanagement über alle Bausteine für eine FAIRe Infrastruktur. Persistente Identifikatoren werden flächendeckend verwendet, die Daten werden in großen internationalen Suchportalen verzeichnet, die FDZ verwenden für die Dokumentation in den meisten Fällen etablierte Metadaten schemata, ermöglichen die Nachnutzung von Metadaten über offene Lizenzen und nutzen zum Teil auch schon semantische Technologien der Wissensrepräsentation. Die Weiterentwicklung der Dateninfrastruktur kann auf diese Metadatenpraktiken und -technologien aufbauen. Dabei ist zu beachten, dass es zwar ein hohes Maß an Professionalität im Metadatenmanagement gibt. Die Praktiken sind aber insgesamt noch stark fragmentiert, was sich unter anderem daran zeigt, dass es keine gemeinsamen technischen Lösungen für das Metadatenmanagement gibt (weder solche, die von den FDZ selbst entwickelt wurden, noch werden kommerzielle Produkte eingesetzt).

Fragmentierung stellt nicht unbedingt ein Problem dar, da sich Metadaten an den Standards und Bedarfen der einzelnen Communities orientieren sollen. Eine verstärkte Harmonisierung der Metadaten ist deshalb zunächst innerhalb von Communities nützlich (so wie dies z. B. im Verbund Forschungsdaten Bildung bereits gehandhabt wird). Eine Harmonisierung der Metadaten über gemeinsam genutzte Werkzeuge (z. B. Metadateneditoren) könnte bei der Verbesserung der Interoperabilität helfen.

Ob noch weitere Maßnahmen für die Verbesserung der FAIRness in den FDZ sinnvoll sind, lässt sich pauschal auf der Grundlage der vorliegenden Erhebung nicht beantworten. Sicher ist, dass größere Investitionen in die Implementierung von Programmierschnittstellen, die Granularität der Datendokumentation und die inhaltliche Erschließung der Daten notwendig wären. Solche Investitionsentscheidungen können nur vor dem Hintergrund konkreter Nutzungsszenarien getroffen werden. Gerade für den automatisierten Zugriff auf (Meta-)Daten sowie die Verknüpfung von Daten fehlen noch konkrete Anforderungen aus den Communities der Nutzenden.

Die Empfehlungen, die aus dem Themenschwerpunkt abgeleitet werden, beziehen sich deshalb vor allem auf solche Aspekte, die es erlauben, mit übersichtlichem Aufwand relevante Lücken in den Informationen über die Daten zu schließen.

³⁸ Betancort Cabrera, N., Bongartz, E. C., Dörrenbächer, N., Goebel, J., Kaluza, H., & Siegers, P. (2020). *White Paper on implementing the FAIR principles for data in the Social, Behavioural, and Economic Sciences*. RatSWD Working Paper 274/2020. Rat für Sozial- und Wirtschaftsdaten (RatSWD). <https://doi.org/10.17620/02671.60>.

8 Beschwerdemanagement

Die Sicherung und Verbesserung der Arbeitsqualität der FDZ gehört zu den Kernaufgaben des RatSWD. Bereits in der Vergangenheit war er Ansprechpartner für Beschwerden, die sich auf das Datenangebot und den Service der FDZ bezogen.

Neben der Durchführung des jährlichen Monitoringverfahrens, dessen Nebenergebnis unter anderem der vorliegende Tätigkeitsbericht ist, hat die Monitoringkommission auch die Aufgabe, ein Anlaufpunkt für Beschwerden von Datennutzenden zu sein.

Um das Beschwerdemanagement zu professionalisieren und transparent zu gestalten, wurde eine Beschwerdestelle bei der Geschäftsstelle des RatSWD eingerichtet. Sie stellt im engen Austausch mit der Monitoringkommission eine schnelle und professionelle Reaktion auf Beschwerden sicher und zieht daraus Anregungen für die Weiterentwicklung der Dateninfrastruktur.

Stellen Datennutzende gravierende Mängel im Angebot eines akkreditierten FDZ fest, so sollten sie diese beim betreffenden FDZ direkt ansprechen und versuchen, eine Lösung zu finden. Wird keine Einigung erzielt, kann das Anliegen an die Beschwerdestelle des RatSWD gerichtet werden.

Die Zuständigkeit der Beschwerdestelle beschränkt sich auf die Prüfung der Einhaltung der Akkreditierungskriterien. Der RatSWD ist nicht Ansprechpartner bei Verzögerungen im normalen Ablauf oder in Bezug auf Verhaltensweisen des Personals der FDZ. Für Beschwerden dieser Art sollte das jeweilige FDZ direkt kontaktiert werden.

Detaillierte Informationen zum Verfahrensweg finden sich im RatSWD Output 8(5)³⁹ bzw. in aktualisierter Form auf der Webseite des RatSWD.⁴⁰

Aktuelle Beschwerdeverfahren des Berichtsjahres 2020

Im Berichtszeitraum 2020 wurde keine Beschwerde an den RatSWD gerichtet.

39 RatSWD [Rat für Sozial- und Wirtschaftsdaten] (2017). *Qualitätssicherung der vom Rat für Sozial- und Wirtschaftsdaten (RatSWD) akkreditierten Forschungsdatenzentren (FDZ)*. RatSWD Output 8(5). <https://doi.org/10.17620/02671.4>.

40 <https://www.konsortswd.de/datenzentren/monitoring-und-beschwerdestelle/beschwerdestelle/> (Zugriff am 31.01.2022).

Anhang

Anhang A: Entwicklung der Forschungsdateninfrastruktur (FDI) des RatSWD48
Anhang B: Index und Angebote der Forschungsdateninfrastruktur (FDI) des RatSWD52
Anhang C: Die Monitoringkommission.....58
Anhang D: Mitwirkende bei der Erstellung des Tätigkeitsberichts 2020.....59



Anhang A

Entwicklung der Forschungsdateninfrastruktur (FDI) des RatSWD

Stand: Dezember 2021

Die FDZ Statistisches Bundesamt, Statistische Ämter der Länder, GML, IZA, BA im IAB und RV sind bereits vor der Gründung des RatSWD eingerichtet worden und gehörten ab 2004 zur Forschungsdateninfrastruktur des RatSWD. Bei diesen FDZ wurde das Jahr der Gründung angegeben. Alle anderen FDZ wurden nach 2004 beim RatSWD akkreditiert. Bei diesen FDZ ist jeweils das Jahr der Akkreditierung angegeben.

Gründung folgender FDZ:



1999

2001

2002

2003

2004

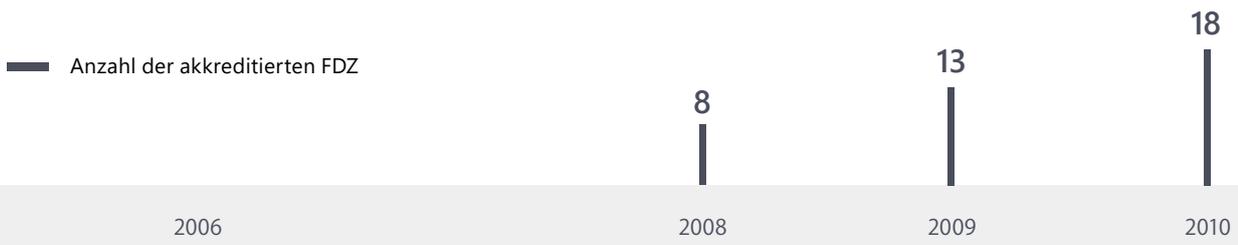
Gründungs-
ausschuss

Empfehlung der „Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik“ (KVI) zur Einrichtung von Forschungsdatenzentren (FDZ)

Konstituierung des Gründungs-
ausschusses des RatSWD)

Gründung
des RatSWD

Akkreditierung folgender FDZ:



Akkreditierung folgender FDZ:



Einführung eines jährliches Monitorings



6. Berufungsperiode

7. Berufungsperiode

Gründung des
KonsortSWD

41 akkreditierte FDZ
im FDI Ausschuss

Anhang B

Index und Angebote der Forschungsdateninfrastruktur (FDI) des RatSWD

Stand: Dezember 2021

Datenangebot: ■ Sozial ■ Wirtschaft ■ Bildung ■ Gesundheit ■ Psychologie ■ Qualitativ ■ Weitere

BAMF-FDZ

(vorläufig
akkreditiert)



Forschungsdatenzentrum des Bundesamtes für Migration und Flüchtlinge

Das BAMF-Forschungsdatenzentrum ist Teil des Forschungszentrums des Bundesamtes. Forschungseinrichtungen können zu Forschungszwecken Zugang zu Daten aus dem Ausländerzentralregister erhalten. Der Datenservice wird in Zukunft durch das Angebot ausgewählter Befragungsdaten erweitert.

<https://www.bamf.de/fdz>



BIBB-FDZ



Forschungsdatenzentrum im Bundesinstitut für Berufsbildung

Personen- und Betriebsdatensätze der Berufsbildungsforschung, die sich inhaltlich hauptsächlich mit dem Erwerb und der Verwertung von beruflichen Kenntnissen und Kompetenzen beschäftigen.

<https://www.bibb.de/de/53.php>



DeZIM.fdz



Forschungsdatenzentrum Deutsches Zentrum für Integration und Migration

Das DeZIM.fdz organisiert den Zugang zu Forschungsdaten, die am Deutschen Zentrum für Integrations- und Migrationsforschung erhoben werden. Zusätzlich bietet das DeZIM.fdz einen umfassenden Support zu diesen Daten und zu verschiedenen methodischen Schwerpunktthemen.

<https://dezim-institut.de/forschungsdatenzentrum-dezimfdz>



EBDC



LMU-ifo Economics & Business Data Center

Datensätze deutscher Unternehmen: Dazu zählen die vom ifo Institut erhobenen Wirtschaftsdaten zur Konjunkturentwicklung, zur Innovationsfreudigkeit und zum Investitionsverhalten sowie externe Daten zur Unternehmensfinanzierung und zur Governance-Struktur von Unternehmen.

<https://www.ifo.de/ebdc>



FDSZ Bundesbank



Forschungsdaten- und Servicezentrum der Bundesbank

Verschiedene Datensätze zu Banken, Wertpapieren, Investmentfonds und Unternehmen sowie Kombinationen daraus; Studie zum Vermögen privater Haushalte.

<https://www.bundesbank.de/de/bundesbank/forschung/fdsz>



FDZ AGD



Forschungsdatenzentrum Archiv für Gesprochenes Deutsch am Institut für Deutsche Sprache

Daten zu gesprochenem Deutsch in Interaktionen (Gesprächskorpora) sowie inländische und ausländische Varietäten des Deutschen (Variationskorpora).

<http://agd.ids-mannheim.de>



FDZ ALLBUS



Forschungsdatenzentrum ALLBUS bei GESIS

Daten aus der allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS) zu Einstellungen, Verhaltensweisen und Sozialstruktur der Bevölkerung in Deutschland.

<https://www.gesis.org/institut/forschungsdatenzentren/fdz-allbus>



FDZ-aviDa (vorläufig akkreditiert)	Forschungsdatenzentrum für audio-visuelle Daten der qualitativen Sozialforschung (FDZ-aviDa) aviDa ist das in Kooperation der Technischen Universität Berlin und der Universität Bayreuth entwickelte Forschungsdatenzentrum (FDZ) für audio-visuelle Daten der empirischen qualitativen Sozialforschung am Fachgebiet für Allgemeine Soziologie der Technischen Universität Berlin. aviDa zielt seit 2018 auf das Öffnen und Teilen videographischer Forschungsdaten. https://fdz-avida.tu-berlin.de/	
FDZ BA im IAB	Forschungsdatenzentrum der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung Daten zu Personen, Haushalten und Betrieben sowie kombinierte Datensätze aus Befragungsdaten und administrativen Daten im Bereich der Sozialversicherung und der Arbeitsmarkt- und Berufsforschung. https://fdz.iab.de	
FDZ-BAuA	Forschungsdatenzentrum der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin Daten aus Quer- und Längsschnittstudien zu Arbeits- und Beschäftigungsbedingungen und deren Auswirkungen auf Gesundheit, Sicherheit und Wohlbefinden von Erwerbstätigen in Deutschland. https://www.baua.de/fdz	
FDZ Bildung	Forschungsdatenzentrum Bildung am DIPF Leibniz-Institut für Bildungsforschung und Bildungsinformation Die angebotenen Daten aus der qualitativen Bildungsforschung umfassen Audio- und Video-Daten, Transkripte und Kontextmaterialien. Im Bereich Erhebungsinstrumente der quantitativen Bildungsforschung werden Testinstrumente, Fragebögen und Skalenkollektionen bereitgestellt. Die Daten und Instrumente entstammen größtenteils der Forschung zu Schul- und Unterrichtsqualität. https://www.fdz-bildung.de	
FDZ-BO	Forschungsdatenzentrum Betriebs- und Organisationsdaten Quantitative und qualitative Betriebs- und Organisationsdaten, verknüpfte Personen- und Betriebsdaten sowie Daten aus Mitarbeitenden- und Mitgliederbefragungen. http://www.fdz-bo.diw.de	
FDZ-Bund	Forschungsdatenzentrum des Statistischen Bundesamtes Flächendeckender Zugang zu amtlichen Mikrodaten aus den Bereichen: Bevölkerung, Bildung, Gesundheit, Wirtschaft, Agrar, Umwelt, Energie, Rechtspflege, Finanzen und Steuern. https://www.forschungsdatenzentrum.de	
FDZ BZgA	Forschungsdatenzentrum der Bundeszentrale für gesundheitliche Aufklärung Daten aus regelmäßig wiederholten bundesweiten Repräsentativbefragungen zur Erreichbarkeit der Bevölkerung mit Aufklärungs- und Präventionsmaßnahmen sowie zum Wissen, den Einstellungen und dem Gesundheitsverhalten der Bevölkerung bezüglich der Gesundheitsthemen der BZgA. https://www.bzga.de/forschung/studien-untersuchungen	
FDZ-Länder	Forschungsdatenzentrum der Statistischen Ämter der Länder Flächendeckender Zugang zu amtlichen Mikrodaten aus den Bereichen: Bevölkerung, Bildung, Gesundheit, Wirtschaft, Agrar, Umwelt, Energie, Rechtspflege, Finanzen und Steuern. https://www.forschungsdatenzentrum.de	

Datenangebot: ■ Sozial ■ Wirtschaft ■ Bildung ■ Gesundheit ■ Psychologie ■ Qualitativ ■ Weitere

FDZ-DJI ■ ■	Forschungsdatenzentrum des Deutschen Jugendinstituts Daten aus den seit 1988 regelmäßig durchgeführten Surveys zum Aufwachsen von Kindern und Jugendlichen sowie zu Lebenslagen von Erwachsenen und Familien. https://www.dji.de/abt2	
FDZ-DZA ■ ■ ■ ■ ■	Forschungsdatenzentrum des Deutschen Zentrums für Altersfragen Daten der Langzeitstudien „Deutscher Alterssurvey (DEAS)“, zum Wandel der Lebenssituationen und Altersverläufe der Bevölkerung in der zweiten Lebenshälfte sowie des Deutschen Freiwilligen Surveys (FWS) - eine repräsentative Bevölkerungsumfrage zum freiwilligen Engagement in Deutschland. https://www.dza.de/forschung/fdz	
FDZ-DZHW ■ ■ ■ ■	Forschungsdatenzentrum für Hochschul- und Wissenschaftsforschung Quantitative und qualitative Daten der Hochschul- und Wissenschaftsforschung, insbesondere Studienberechtigtenpanel, Absolventenpanel, Sozialerhebung und Wissenschaftsbefragung des Deutschen Zentrums für Hochschul- und Wissenschaftsforschung (DZHW). https://fdz.dzhw.eu	
FDZ eLabour ■ ■ ■	Forschungsdatenzentrum eLabour Qualitative Daten aus Studien zur Arbeits- und Industriesoziologie mit dem Schwerpunkt Wandel von Arbeit, u. a. offene bis halbstandardisierte Interviews, Beobachtungen sowie Expertengespräche. http://elabour.de	
FDZ GePaRD <i>(vorläufig akkreditiert)</i> ■	Pharmakoepidemiologische Forschungsdatenbank Das FDZ GePaRD basiert auf Daten gesetzlicher Krankenversicherungen in Deutschland seit 2004. Mit GePaRD können u. a. Forschungsfragen zum Gebrauch und zur Sicherheit von Arzneimitteln und Impfstoffen in der Routineversorgung untersucht werden, sofern die jeweilige Datennutzung gemäß § 75 SGB X genehmigt wird. https://www.bips-institut.de/forschung/forschungsinfrastrukturen/gepard.html	
FDZ GML ■ ■ ■	Forschungsdatenzentrum German Microdata Lab bei GESIS Forschungsbasierter Service zu Mikrodaten der amtlichen Statistik in Europa und Deutschland: Tools für Datenmanagement und Datenanalyse. Metadaten (MISSY): Umfassende und detaillierte Datendokumentation amtlicher Mikrodaten. Wissenstransfer: Beratung, Workshops und Nutzerkonferenzen zu methodischen und inhaltlichen Forschungsfragen bei der Analyse amtlicher Mikrodaten. Gegründet 1987. https://www.gesis.org/institut/forschungsdatenzentren/fdz-german-microdata-lab	
FDZ Internationale Umfrageprogramme ■ ■ ■ ■ ■ ■	Forschungsdatenzentrum Internationale Umfrageprogramme bei GESIS International vergleichende Umfragedaten aus über 70 Ländern der Welt zu fast allen Themen der Sozialwissenschaften: Comparative Study of Electoral Systems (CSES), European Values Study (EVS), Eurobarometer, Europa-Wahlen (EES), International Social Survey Programme (ISSP). https://www.gesis.org/institut/forschungsdatenzentren/fdz-internationale-umfrageprogramme	

FDZ IQB 	Forschungsdatenzentrum am Institut zur Qualitätsentwicklung im Bildungswesen <p>Deutsche Datensätze aus den großen nationalen und internationalen Schulleistungsstudien sowie aus nationalen Studien mit Kompetenzmessungen im Bildungsbereich.</p> <p>https://www.iqb.hu-berlin.de/fdz</p>	
FDZ-IWH 	Forschungsdatenzentrum des Leibniz-Instituts für Wirtschaftsforschung Halle <p>Unternehmensdaten zu Entwicklungstendenzen in der Industrie und im Baugewerbe Ostdeutschlands, zu Privatisierungsaktivitäten der Treuhandanstalt, zur Standortwahl multinationaler Unternehmen in den neuen Bundesländern und in Mittel- und Osteuropa sowie zu Produktivitäts- und Wettbewerbsfähigkeitsindikatoren europäischer Länder nach Firmengrößeklassen.</p> <p>https://www.iwh-halle.de/forschung/daten-und-analysen</p>	
FDZ IZA, IDSC 	Internationales Forschungsdatenzentrum des Forschungsinstituts zur Zukunft der Arbeit <p>Nationale und internationale Arbeitsmarktdatensätze mit standardisierten Informationen (eddi-conferences.eu). Forschung sowie Methoden und Ressourcen für die Nutzung von Onlinedaten in Arbeitsmarktökonomie und Sozialwissenschaften. Entwicklung von Werkzeugen und Methoden für Remote Access (statsdirect.org) und Remote Processing (JoSuA).</p> <p>https://www.iza.org/forschung/idsc</p>	
FDZ im KBA 	Forschungsdatenzentrum im Kraftfahrt-Bundesamt <p>Mit den vom KBA geführten Registern als Datengrundlage, umfasst das Datenangebot des FDZ im KBA derzeit quantitative Mikrodaten zum Zugang in das Fahrereignisregister.</p> <p>https://www.kba.de/DE/Statistik/Forschungsdatenzentrum/forschungsdatenzentrum_node.html</p>	
FDZ-LifBi 	Forschungsdatenzentrum des Leibniz-Instituts für Bildungsverläufe e.V. an der Otto-Friedrich-Universität Bamberg <p>Längsschnittdaten des seit 2010 laufenden Nationalen Bildungspanels (NEPS - National Educational Panel Study) mit über 60.000 Teilnehmenden in sechs Startkohorten zur Erforschung von Kompetenzentwicklungen, Bildungsprozessen, Bildungsentscheidungen und Bildungsrenditen in formalen, nicht-formalen und informellen Kontexten.</p> <p>https://www.lifbi.de/Institut/Organisation/Forschungsdatenzentrum-LifBi</p>	
FDZ-pairfam 	Forschungsdatenzentrum des Beziehungs- und Familienpanels <p>Datensätze aus der repräsentativen, interdisziplinären Längsschnittstudie zu partnerschaftlichen und familialen Lebensformen in Deutschland: „Panel Analysis of Intimate Relationships and Family Dynamics“ (pairfam).</p> <p>https://www.pairfam.de</p>	
FDZ PIAAC 	Forschungsdatenzentrum PIAAC bei GESIS <p>Nationale und internationale Datensätze zu PIAAC (Programme for the International Assessment of Adult Competencies). Für Deutschland stehen des Weiteren Regionaldaten sowie Längsschnittdaten zur Verfügung.</p> <p>https://www.gesis.org/institut/forschungsdatenzentren/fdz-piaac</p>	

Datenangebot: ■ Sozial ■ Wirtschaft ■ Bildung ■ Gesundheit ■ Psychologie ■ Qualitativ ■ Weitere

FDZ PsychData des ZPID ■ ■ ■	Forschungsdatenzentrum PsychData des Leibniz-Zentrums für Psychologische Information und Dokumentation Bündelung quantitativer Datensätze sowohl aus der Grundlagenforschung als auch aus der angewandten Psychologie; Ausrichtung des Sammelauftrages insbesondere auf Längsschnittstudien, groß angelegte Umfragestudien und Testentwicklungen. https://www.psychdata.de	 FDZ am ZPID
FDZ Qualiservice ■ ■ ■ ■ ■	Forschungsdatenzentrum Qualiservice Qualiservice bietet die Archivierung, Kuration und Bereitstellung qualitativer Forschungsdaten unterschiedlicher Disziplinen an. Die Services sind sicher, flexibel und forschungsorientiert und beinhalten umfassende Beratung und Aufbereitung von qualitativen Primärstudien für die Nachnutzung, Langzeitarchivierung sowie die Bereitstellung archivierter Forschungsdaten und relevanter Kontextinformationen. https://www.qualiservice.org	
FDZ RKI ■ ■	Forschungsdatenzentrum des Robert Koch-Instituts Daten zum Gesundheitszustand und zum Gesundheitsverhalten der Bevölkerung, die auf der Basis von bundesweit repräsentativen Studien gewonnen wurden. https://www.rki.de/puf	
FDZ Ruhr am RWI ■ ■ ■	Forschungsdatenzentrum Ruhr am RWI – Leibniz-Institut für Wirtschaftsforschung Spezialisierung auf regionale Daten: Angebot von Daten auf Ein-Quadratkilometer-Rastern. Neben der Geo-Referenzierung von Daten auf wissenschaftlicher Basis werden unterschiedliche Individual- und Unternehmensdaten aus Forschungsprojekten des RWI zur Verfügung gestellt. https://www.rwi-essen.de/forschung-und-beratung/fdz-ruhr	
FDZ-RV ■ ■ ■ ■	Forschungsdatenzentrum der Rentenversicherung Daten zu Versicherungskonten der in der gesetzlichen Rentenversicherung Versicherten, in denen der Versicherungsverlauf sowie Leistungen an die Versicherten aus Rente und Rehabilitation gespeichert sind. http://www.fdz-rv.de/FdzPortalWeb	
FDZ-SHARE ■ ■ ■ ■ ■ ■	Forschungsdatenzentrum des Survey of Health, Ageing and Retirement in Europe Daten der multidisziplinären Längsschnittstudie „SHARE“, die in bislang sieben Wellen Daten von ca. 140.000 Individuen zu Gesundheit, sozio-ökonomischem Status sowie sozialen und familiären Netzwerken von Personen ab 50 Jahren in über 20 europäischen Ländern und in Israel erhoben hat. Die 8. Welle von SHARE wurde 2020 erhoben. http://www.share-project.org/data-access.html	
FDZ SOEP ■ ■ ■ ■ ■ ■	Forschungsdatenzentrum des Sozio-oekonomischen Panels am DIW Berlin Daten aus repräsentativ erhobenen Wiederholungsbefragungen von Privathaushalten. Die SOEP-CORE Stichprobe umfasst Themen wie Einkommen, Erwerbstätigkeit, Bildung und Gesundheit. Zudem gibt es die längsschnittliche Innovationsstichprobe (SOEP-IS), die es externen Forschenden ermöglicht, eigene Forschungsvorhaben einzubringen. https://www.diw.de/SOEPfdz	

FDZ Wahlen	Forschungsdatenzentrum Wahlen bei GESIS	
■ ■	<p>Zugang zu nationalen Wahlstudien (Bundes- und Landtagswahlen), Politbarometer, Forsa-Bus, ARD Deutschlandtrend, Umfragen der Bundesregierung. Das derzeit größte Projekt im FDZ Wahlen ist die German Longitudinal Election Study (GLES).</p> <p>https://www.gesis.org/institut/forschungsdatenzentren/fdz-wahlen</p>	
FDZ Wissenschaftsstatistik	Forschungsdatenzentrum Wissenschaftsstatistik des Stifterverbandes	
■ ■ ■	<p>Daten über die Forschungs- und Entwicklungsaktivitäten der deutschen Wirtschaft, zum finanziellen Volumen, der Struktur und der regionalen Verteilung von Forschung und Entwicklung (FuE) sowie der in diesem Bereich tätigen Personen.</p> <p>https://www.fdz-wissenschaftsstatistik.de</p>	
IÖR-Monitor	Monitor der Siedlungs- und Freiraumentwicklung	
■	<p>Der IÖR-Monitor ist eine Forschungsdateninfrastruktur des Leibniz-Instituts für ökologische Raumentwicklung (IÖR). Er stellt Daten und Informationen in hoher räumlicher Auflösung zur Flächennutzungsstruktur und deren Entwicklung sowie zur Landschaftsqualität für die Bundesrepublik Deutschland bereit.</p> <p>https://www.ioer-monitor.de</p>	
SAFE FDZ <i>(vorläufig akkreditiert)</i>	Forschungsdatenzentrum des Leibniz-Instituts für Finanzmarktforschung SAFE	
■ ■	<p>Der Mangel an pan-europäischen Finanzdaten führt dazu, dass Forscher auf US Daten ausweichen und Forschungsergebnisse nur eingeschränkt auf den europäischen Raum übertragen können. Das SAFE Forschungsdatenzentrum bündelt nicht nur vorhandene Daten, sondern erhebt und erstellt auch neue deutsche und europäische Datensätze, um die europäische Perspektive der empirischen Forschung zu stärken.</p> <p>https://safe-frankfurt.de/de/data-center.html</p>	
ZEW-FDZ	Forschungsdatenzentrum des Zentrums für Europäische Wirtschaftsforschung	
■ ■	<p>Das ZEW-FDZ stellt die Individualdaten von ZEW-Firmenbefragungen zu Innovationsaktivitäten, zur Entwicklung junger Unternehmen, zur Nutzung von Informations- und Kommunikationstechnologien und anderen Themen zur Verfügung. Ferner können die Daten einer ZEW-Expertenbefragung, der „ZEW-Finanzmarkttest“, genutzt werden.</p> <p>https://www.zew.de/zew-fdz</p>	

Anhang C

Die Monitoringkommission

Für die Qualitätssicherung hat der RatSWD die Etablierung einer Monitoringkommission ab Juli 2016 beschlossen. Ihre Hauptaufgabe ist das Sammeln und Bewerten der regelmäßigen Berichte der FDZ. Ferner überprüft die Kommission die Erfüllung der Auflagen im Fall von vorläufigen Akkreditierungen. Die Kommission wird für je drei Jahre zeitlich parallel zur RatSWD-Legislaturperiode vom FDI Ausschuss aus den eigenen Reihen gewählt und genießt damit besonderes Vertrauen und Legitimität. Sie besteht aus vier Mitgliedern des FDI Ausschusses und zwei Stellvertretenden (für den Fall der Verhinderung eines gewählten Mitglieds) sowie den Vorsitzenden des RatSWD als Gästen.

Mitglieder der Monitoringkommission

Lea Eilers (bis Juli 2021)

Forschungsdatenzentrum Ruhr am RWI – Leibniz-Institut für Wirtschaftsforschung (FDZ Ruhr am RWI)

Dr. Benjamin Fuchs

Forschungsdatenzentrum im Kraftfahrt-Bundesamt (FDZ im KBA)

Dr. Cornelia Lang

Forschungsdatenzentrum des Leibniz-Instituts für Wirtschaftsforschung Halle (FDZ-IWH)

Dr. Laura Menze (seit August 2021)

Forschungsdatenzentrum der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (FDZ-BAuA)

Dana Müller (Vorsitz)

Forschungsdatenzentrum der Bundesagentur für Arbeit (BA) im Institut für Arbeitsmarkt- und Berufsforschung (FDZ BA im IAB)

Holger Quellenberg

Forschungsdatenzentrum des Deutschen Jugendinstituts (FDZ-DJI)

Dr. Pascal Siegers

Forschungsdatenzentrum ALLBUS bei GESIS (FDZ ALLBUS)

Ständige Gäste der Monitoringkommission

Prof. Dr. Monika Jungbauer-Gans

Vorsitzende des RatSWD (seit Juli 2020)

Prof. Dr. Kerstin Schneider

Stellvertretende Vorsitzende des RatSWD (seit Juli 2020)

Anhang D

Mitwirkende bei der Erstellung des Tätigkeitsberichts 2020

Florence Baillet

Geschäftsstelle des Rates für Sozial- und Wirtschaftsdaten (RatSWD)

Dr. Benjamin Fuchs

Forschungsdatenzentrum im Kraftfahrt-Bundesamt (FDZ im KBA)

Dr. Anna Fräßdorf

Geschäftsstelle des Rates für Sozial- und Wirtschaftsdaten (RatSWD)

Dr. Cornelia Lang

Forschungsdatenzentrum des Leibniz-Instituts für Wirtschaftsforschung Halle (FDZ-IWH)

Dr. Laura Menze

Forschungsdatenzentrum der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (FDZ-BAuA)

Dana Müller

Forschungsdatenzentrum der Bundesagentur für Arbeit (BA) im Institut für Arbeitsmarkt- und Berufsforschung (FDZ BA im IAB)

Holger Quellenberg

Forschungsdatenzentrum des Deutschen Jugendinstituts (FDZ-DJI)

Thomas Runge

Geschäftsstelle des Rates für Sozial- und Wirtschaftsdaten (RatSWD)

Lea Salathé

Geschäftsstelle des Rates für Sozial- und Wirtschaftsdaten (RatSWD)

Dr. Pascal Siegers

Forschungsdatenzentrum ALLBUS bei GESIS (FDZ ALLBUS)

Impressum

Herausgeber:

Rat für Sozial- und Wirtschaftsdaten (RatSWD)
Geschäftsstelle
Am Friedrichshain 22
10407 Berlin
office@ratswd.de
<https://www.ratswd.de>

Redaktion:

Florence Baillet, Dr. Anna Fräßdorf, Lea Salathé

Gestaltung/Satz:

Claudia Kreutz

Icons:

made by Freepik from <https://www.flaticon.com>

Berlin, März 2022

Die Geschäftsstelle des RatSWD wird als Teil von KonsortSWD im Rahmen der NFDI durch die Deutsche Forschungsgemeinschaft (DFG) gefördert - Projektnummer: 442494171.



Diese Veröffentlichung ist unter der Creative-Commons-Lizenz (CC BY 4.0) lizenziert:

<https://creativecommons.org/licenses/by/4.0/>

Ausgenommen von der oben genannten Lizenz sind Teile, Abbildungen und sonstiges Drittmaterial, wenn anders gekennzeichnet, sowie die Coverbilder auf Seite 1 und 64.

doi: 10.17620/02671.65

Zitationsvorschlag:

RatSWD [Rat für Sozial- und Wirtschaftsdaten] (2022). *Tätigkeitsbericht 2020 der vom RatSWD akkreditierten Forschungsdatenzentren (FDZ)*. Berlin. Rat für Sozial- und Wirtschaftsdaten (RatSWD). <https://doi.org/10.17620/02671.65>.

■ **Der Rat für Sozial- und Wirtschaftsdaten (RatSWD)** berät seit 2004 die Bundesregierung und die Regierungen der Länder in Fragen der Forschungsdateninfrastruktur für die empirischen Sozial-, Verhaltens- und Wirtschaftswissenschaften. Im RatSWD arbeiten zehn durch Wahl legitimierte Vertreterinnen und Vertreter der sozial-, verhaltens- und wirtschaftswissenschaftlichen Fachdisziplinen mit zehn Vertreterinnen und Vertretern der wichtigsten Datenproduzenten zusammen.

Der RatSWD ist Teil des Konsortiums für die Sozial-, Verhaltens-, Bildungs- und Wirtschaftswissenschaften (KonsortSWD) in der Nationalen Forschungsdateninfrastruktur (NFDI). Er versteht sich als institutionalisiertes Forum des Dialoges zwischen Wissenschaft und Datenproduzenten und erarbeitet Empfehlungen und Stellungnahmen. Dabei engagiert er sich für eine Infrastruktur, die der Wissenschaft einen breiten, flexiblen und sicheren Datenzugang ermöglicht. Diese Daten werden von staatlichen, wissenschaftsgetragenen und privatwirtschaftlichen Akteuren bereitgestellt. Derzeit hat der RatSWD 41 Forschungsdatenzentren (Stand: Februar 2022) akkreditiert und fördert deren Kooperation.

