

# 20 Jahre Forschungsdatenzentrum der Statistischen Ämter der Länder

## Das Daten- und Dienstleistungsangebot für wissenschaftliche Nutzungen von Mikrodaten der amtlichen Statistik

Dipl.Soz.Univ. Anja Hlawatsch, Karen Meyer, M.Sc. und Dipl.Soz.Wiss.Univ. Patrick Rothe

Das Forschungsdatenzentrum (FDZ) der Statistischen Ämter der Länder hat es sich seit seiner Gründung im April 2002 zum Ziel gesetzt, den Zugang zu Mikrodaten der amtlichen Statistik für wissenschaftliche Forschungszwecke zu ermöglichen. Dieser wurde seitdem fortlaufend verbessert und an die Bedarfe der Wissenschaft angepasst, sodass heute auf eine Vielzahl hochwertiger Daten zurückgegriffen werden kann. Offene Fragestellungen im Bereich von Politik und Wirtschaft machen komplexe Forschungen notwendig, welche nur mithilfe der Mikrodaten der amtlichen Statistik durchgeführt werden können. Als Serviceeinrichtung der Statistischen Ämter der Länder ist das FDZ für viele Wissenschaftlerinnen und Wissenschaftler im Laufe seines 20-jährigen Bestehens zu einer unverzichtbaren Institution geworden.

Dieser Beitrag nimmt das 20-jährige Jubiläum zum Anlass, das FDZ und seine Datenangebote vorzustellen sowie einen Blick in die Vergangenheit zu werfen. Abschließend wird speziell auf den Standort Bayern eingegangen und ein kurzer Ausblick auf zukünftige Herausforderungen gegeben.

### Was ist das FDZ?

Während noch in den 1990er-Jahren der Zugang zu Mikrodaten in Deutschland für Wissenschaftlerinnen und Wissenschaftler nur unter großem Aufwand möglich war, hat sich mittlerweile das FDZ als wichtige Serviceeinrichtung für die Bereitstellung von amtlichen Mikrodaten etabliert. Grundlage für die Gründung der FDZ war das Gutachten der „Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik“ (KVI), welches 1999 im Auftrag des Bundesministeriums für Bildung und Forschung (BMBF) verschiedene Vorschläge zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik erarbeitet hat. Zwei zentrale Ergebnisse waren die Einrichtung eines Rates für Sozial- und Wirtschaftsdaten (RatSWD) sowie der Aufbau von Forschungsdatenzentren.<sup>1</sup> Im Herbst 2001 wurde zunächst das FDZ des Statistischen Bundesamtes gegründet. Wenig später,

im April 2002, entstand das FDZ der Statistischen Ämter der Länder als gemeinsames Pilotprojekt aller Statistischen Ämter der Länder. Beide FDZ sind zwar organisatorisch voneinander getrennte Einrichtungen, dennoch arbeiten sie im Rahmen einer Arbeitsgemeinschaft eng zusammen und bieten gemeinsam ein abgestimmtes Daten- und Dienstleistungsangebot für wissenschaftliche Nutzungen an. Die Kooperationsvereinbarung wurde im Februar 2003 durch den damaligen Präsidenten des Statistischen Bundesamtes, Johann Hahlen, und den Präsidenten des Bayerischen Landesamts für Statistik und Datenverarbeitung, Dr. Peter Bauer, in seiner Funktion als Vorsitzender des Lenkungsausschusses des FDZ der Statistischen Ämter der Länder unterzeichnet. So gilt in beiden FDZ das Selbstverständnis, als zentrales Bindeglied zwischen amtlicher Statistik und der Wissenschaft zu agieren. Neben drei Standorten des FDZ des Bundes gibt es aktuell 14 regionale

<sup>1</sup> Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (KVI) (2001): Wege zu einer besseren informationellen Infrastruktur. Baden-Baden: Nomos.



### RatSWD – Rat für Sozial- und Wirtschaftsdaten

- unabhängiges Gremium von empirisch arbeitenden Forscherinnen und Forschern sowie von Vertreterinnen und Vertretern der wichtigsten öffentlichen Einrichtungen zur Datenerhebung
- Gründung: 2004 als Beirat der Bundesregierung
- Koordination von aktuell 41 akkreditierten Datenzentren
- Zentrale Aufgaben:
  - Weiterentwicklung der Forschungsdateninfrastruktur
  - Interessenvertretung von Datenproduzierenden und -nutzenden
  - Beratungsfunktion für Politik und Wissenschaftsinstitutionen
  - Akkreditierung und Evaluation der Arbeit von Forschungsdatenzentren
  - Europäische und internationale Vernetzung von Forschungsdateninfrastrukturen
  - Austausch durch Dialogveranstaltungen, Fachtagungen und die Konferenz für Sozial- und Wirtschaftsdaten (KSWD)

2 KonsortSWD: Auf einen Blick: [www.konsortswd.de/ratswd/der-ratswd/auf-einen-blick/](http://www.konsortswd.de/ratswd/der-ratswd/auf-einen-blick/), abgerufen am 29.09.2022

3 KonsortSWD: Akkreditierungsverfahren und -kriterien: [www.konsortswd.de/datenzentren/akkreditierung/](http://www.konsortswd.de/datenzentren/akkreditierung/), abgerufen am 29.09.2022.

4 RatSWD (2011): Auf Erfolge aufbauend – Zur Weiterentwicklung der Forschungsinfrastruktur für die Sozial- Verhaltens- und Wirtschaftswissenschaften. Empfehlungen des Rates für Sozial- und Wirtschaftsdaten. Opladen: Budrich.

5 Wissenschaftsrat (2011): Empfehlungen zu Forschungsdateninfrastrukturen in den Geistes- und Sozialwissenschaften. Berlin.

Standorte des FDZ der Länder – und zusätzlich eine Reihe von Außenstellen in Hochschulen und sonstigen wissenschaftlichen Einrichtungen innerhalb Deutschlands. Eine bei Information und Technik Nordrhein-Westfalen angesiedelte Geschäftsstelle übernimmt die Außenvertretung des FDZ der Statistischen Ämter der Länder, die Bearbeitung von Grundsatzfragen sowie die Koordination und Verwaltung der regelmäßig anfallenden Arbeiten des FDZ.

Das FDZ des Statistischen Bundesamtes und das FDZ der Statistischen Ämter der Länder gehören zu den ersten Institutionen, die vom RatSWD<sup>2</sup> (siehe Infokasten „RatSWD – Rat für Sozial und Wirtschaftsdaten“) formal als Forschungsdatenzentren akkreditiert wurden. Seitdem sind weitere Institutionen in Wissenschaft und Verwaltung

dazu übergegangen, der Wissenschaft einen flexiblen und umfangreichen Datenzugang zu ermöglichen. In den Sozial-, Verhaltens- und Wirtschaftswissenschaften hat sich erfolgreich eine Forschungsdateninfrastruktur (FDI) in Form von Forschungsdatenzentren etabliert. Um offiziell als FDZ anerkannt zu werden, müssen Mindeststandards und Akkreditierungskriterien erfüllt werden, welche vom RatSWD und dem ständigen Ausschuss für Forschungsdateninfrastruktur (FDI Ausschuss) erarbeitet wurden.<sup>3</sup>

### Finanzierung

Seit 2011 kann das FDZ als dauerhaft etablierte Einrichtung angesehen werden. Zuvor wurde das FDZ zunächst im Rahmen eines Pilotprojekts und in einer ersten Förderphase von 2004 bis 2007 vom BMBF finanziert. Im Fokus standen damals die Einrichtung und der Betrieb einer fachlich zentralisierten Datenhaltung, die Entwicklung und Pflege eines Metadateninformationssystems sowie die Bereitstellung von Gastwissenschaftsarbeitsplätzen (GWAP) und Kontrollierter Datenfernverarbeitung (KDFV). Nach überaus positivem Votum startete im Jahr 2007 die zweite Förderphase des BMBF bis 2010, während der die Weiterentwicklung des FDZ erfolgreich vorangetrieben werden konnte. Aufgrund der nachdrücklichen Aufforderung durch den RatSWD<sup>4</sup> und den Wissenschaftsrat<sup>5</sup>, das FDZ als dauerhaften Bestandteil der amtlichen Statistik zu etablieren, wird das FDZ heute finanziell von den Statistischen Landesämtern und zu einem Teil durch die Entgelte der Nutzerinnen und Nutzer getragen.

### Fachlich zentralisierte Datenhaltung

In Deutschland werden über 90 % aller amtlichen Statistiken dezentral in den einzelnen Statistischen Ämtern der Länder erhoben, aufbereitet und gespeichert. Da sich der Fokus wissenschaftlicher Forschungsprojekte in der Regel jedoch auf mehrere Länder oder sogar das gesamte Bundesgebiet richtet, haben die FDZ der Statistischen Ämter des Bundes und der Länder eine fachlich zentralisierte Datenhaltung etabliert. So sind für alle angebotenen Statistiken länderübergreifende Datensätze für das gesamte Bundesgebiet verfügbar. Im Rahmen der zentralisierten Datenhaltung

haben die FDZ-Standorte – in Anlehnung an das Patenlandprinzip der amtlichen Statistik – jeweils die fachliche Zuständigkeit für bestimmte Statistiken übernommen. Der fachlich zuständige Standort übernimmt die Aufbereitung der FDZ-Daten und steht den Nutzerinnen und Nutzern als fachlicher Ansprechpartner für Fragen hinsichtlich der jeweiligen Statistik beratend zur Seite, ohne jedoch dabei die Datenhoheit der anderen Landesämter zu beschränken.

### Notwendigkeit von Mikrodaten

Grundsätzlich wird zwischen Mikrodaten (auch Einzel- bzw. Individualdaten genannt) und Aggregatdaten unterschieden. Während es sich bei Mikrodaten um Angaben zu einzelnen Einheiten einer Erhebung, beispielsweise Personen, Haushalten oder Unternehmen handelt, beschreibt der Begriff Aggregatdaten zusammengefasste Daten, die bereits statistische Ergebnisse darstellen. Mikrodaten beinhalten folglich deutlich mehr Detailinformationen als Aggregatdaten und bieten demnach ein entsprechend größeres Analysepotenzial. Die amtlichen Statistiken beruhen häufig auf sehr großen Beobachtungszahlen und sind oftmals sogar Vollerhebungen. Für wissenschaftliche Forschungsprojekte sind die daraus resultierenden detaillierten Analysemöglichkeiten von zentraler Bedeutung. Verfahren der Inferenzstatistik, wie Hypothesentests, und multivariate Verfahren, wie Regressionsanalysen, erfordern detaillierte Individualdaten.

### „Kleine FDZ-Statistik“

Die FDZ der Statistischen Ämter des Bundes und der Länder bieten mittlerweile Gastwissenschafts-arbeitsplätze in 20 über ganz Deutschland verteilten Städten an. Diese breite regionale Streuung ermöglicht den Forscherinnen und Forschern einen arbeitsstätten- beziehungsweise heimatnahen Datenzugang und hält somit die Aufwände hinsichtlich Reisekosten und Zeitaufwand möglichst gering.

Über 6 300 Personen aus mehr als 400 inländischen, aber auch ausländischen wissenschaftlichen Einrichtungen haben in fast 4 000 wissenschaftlichen Projekten seit Einrichtung der

FDZ-Standorte mit den Mikrodaten der amtlichen Statistik geforscht, was die enorme Reichweite der FDZ deutlich macht.

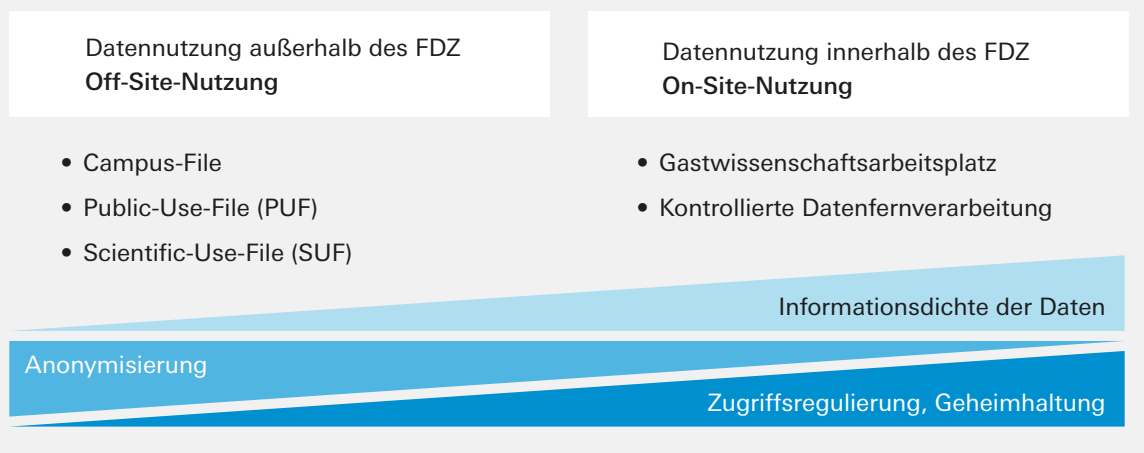
Das umfangreiche Portfolio von mittlerweile über 2 000 Produkten aus 129 Statistiken erreicht Forscherinnen und Forscher verschiedenster Disziplinen. Insbesondere die Sozial-, Wirtschafts- und Gesundheitsstatistiken werden sehr häufig nachgefragt. Das Angebotsspektrum umfasst auch Statistiken aus den Bereichen Agrar, Bauen und Wohnen, Bildung, Energie, Recht und Umwelt, welche ebenfalls intensiv genutzt werden. Die nutzungsstärksten Statistiken sind auf Produktebene der Mikrozensus, das AFiD-Panel Industriebetriebe (ein eigens von den FDZ entwickeltes Datenprodukt) und die DRG-Statistik, eine jährliche Vollerhebung aller nach Fallpauschalen abgerechneten vollstationären Krankenhaufälle innerhalb Deutschlands. Die breit gefächerten FDZ-Datenbestände sind somit eine wichtige Grundlage für die empirische Erforschung und wissenschaftliche Bearbeitung vieler der aktuellen „Megathemen“ unserer Zeit.

### Zugangswege

Die Arbeit an den Mikrodaten im FDZ ist über verschiedene Zugangswege möglich, die sich auch hinsichtlich des Anonymisierungsgrads beziehungsweise – daraus resultierend – des Informationsgehalts des Datenmaterials unterscheiden (siehe Abb. 1).

Zunächst sind die Off-Site-Nutzungen zu nennen. Die hierunter fallenden CAMPUS-Files (CFs), Public-Use-Files (PUFs) und Scientific-Use-Files (SUFs) dürfen auch außerhalb der geschützten Räumlichkeiten der amtlichen Statistik verwendet werden. Im Falle der CFs und der PUFs ist nach § 16 Absatz 1 Satz 2 Punkt 4 des Bundesstatistikgesetzes (BStatG) gar eine Nutzung seitens der breiten Öffentlichkeit möglich. Der Anonymisierungsgrad ist bei diesen Materialien entsprechend sehr hoch (absolute Anonymität), woraus sich im Gegenzug jedoch ein relativ geringer Informationsgehalt der Daten ergibt. Dem Nutzungsziel von CFs und PUFs, nämlich dem Sammeln von Erfahrungen in der Auswertung von Mikrodaten im

Abb. 1  
Zugangswege für Mikrodaten der Statistischen Ämter



Lehrbereich, steht dies allerdings nicht entgegen. Im Unterschied zu CFs und PUFs können SUFs ausschließlich innerhalb der Räumlichkeiten von wissenschaftlichen Einrichtungen genutzt werden, sofern diese im Bundesgebiet liegen. Durch ein dementsprechend schwächeres Anonymisierungslevel eröffnen sie dafür die Möglichkeit, aussagekräftige inhaltliche Analysen durchzuführen und sind daher sehr beliebt.

Neben den Off-Site-Nutzungen existieren mit den Gastwissenschaftsarbetsplätzen (GWAPs) und der Kontrollierten Datenfernverarbeitung (KDFV) zwei weitere Zugangswege: die sogenannten On-Site-Nutzungen. Das Kernelement ist hierbei, dass das Mikrodatenmaterial die Räumlichkeiten der amtlichen Statistik nicht verlässt. Die GWAPs sind abgeschottete Arbeitsplätze der amtlichen Statistik, an denen die Nutzerinnen und Nutzer die Einzeldaten analysieren können. Den Forscherinnen und Forschern stehen hier Computerarbeitsplätze mit den gängigsten wissenschaftlichen Analyseprogrammen (STATA, R, SAS und SPSS) zur Verfügung. Die generierten Ergebnisse werden von Seiten des FDZ-Personals nach Erstellung auf die Einhaltung der statistischen Geheimhaltungsvorschriften geprüft und erst dann an die Nutzerinnen und Nutzer verschickt. Nutzungsberechtigt sind nur Angehörige von Einrichtungen mit der Aufgabe unabhängiger wissenschaftlicher Forschung. Aufgrund des Schutzes der Einzeldaten durch die Regulierung

des Datenzugangs und die Ausstattung des PC-Arbeitsplatzes können am GWAP – je nach Sensibilität der Daten – faktisch oder formal anonyme Mikrodaten bereitgestellt werden. Voraussetzung für die Nutzung des GWAPs ist jedoch – wie auch bei der Nutzung von SUFs – immer eine Verpflichtung der Datennutzerinnen und -nutzer auf die Wahrung der statistischen Geheimhaltung nach §16 Abs.7 BStatG.

Wird die kontrollierte Datenfernverarbeitung (KDFV) genutzt, erhalten die Forscherinnen und Forscher keinen direkten Zugang zu den Mikrodaten. Stattdessen werden Dummy-Files – sogenannte Datenstrukturfiles (DSF) – zur Verfügung gestellt, die in Aufbau und Struktur, nicht jedoch hinsichtlich des Inhalts, dem Originalmaterial entsprechen. Anhand der DSF können die Nutzerinnen und Nutzer ihre Programmcodes in STATA, R, SAS oder SPSS entwickeln. Mittels dieser Codes kann das FDZ-Personal dann die Originaldaten auswerten. Der Vorteil dieses Nutzungswegs ist die Möglichkeit für die Nutzerinnen und Nutzer, auch Daten auswerten zu können, die in dieser Detailtiefe nicht zur Arbeit mit direkter Sicht auf die Daten bereitgestellt werden könnten.

Die entstandenen Ergebnisse werden, wie auch im Fall der GWAPs, durch das FDZ-Personal hinsichtlich der statistischen Geheimhaltung geprüft und den Nutzerinnen und Nutzern zur Verfügung gestellt (siehe Infokasten „Statistische Geheimhaltung



### Statistische Geheimhaltung und Anonymisierung

Der Zugang zu amtlichen Mikrodaten steht seit jeher in einem Spannungsverhältnis zwischen Datenschutz und Wissenschaftsfreiheit. § 16 Abs. 6 des BStatG stellt die rechtliche Grundlage für das Datenangebot der FDZ dar. Hier werden die rechtlichen Rahmenbedingungen für die Weitergabe statistischer Einzelangaben an Forschungseinrichtungen geregelt. Eine Übermittlung der Daten ist nur möglich, wenn diese den Anforderungen faktischer Anonymität genügen. Dies bedeutet, dass die Einzelangaben nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft den jeweils dahinter stehenden Merkmalsträgern, beispielsweise Einzelpersonen, zugeordnet werden können. Eine Nutzung von formal anonymen Daten ist nur in Räumlichkeiten der amtlichen Statistik zulässig, wo bestimmte technische und organisatorische Maßnahmen dafür sorgen, dass ein ausreichender Schutz der Daten gewährleistet werden kann.

In ihrer Mittlerfunktion zwischen Datenanbietern und Wissenschaft müssen die FDZ sowohl die Rechte der Forscherinnen und Forscher, als auch der Auskunftspflichtigen wahren. Neben dem Recht der Freiheit von Wissenschaft und Forschung (Art. 5 Abs. 1 GG) auf der einen Seite steht das Recht auf informationelle Selbstbestimmung (Art. 2 Abs. 1 GG) der einzelnen Bürgerinnen und Bürger auf der anderen Seite. Dazu kommt das Prinzip der statistischen Geheimhaltung. Um das Vertrauen, welches die Befragten und Auskunftspflichtigen der amtlichen Statistik entgegenbringen, nicht zu schädigen, steht hierbei die Wahrung des Statistikgeheimnisses im Fokus. So muss durch die FDZ verpflichtend sichergestellt werden, dass – analog zu allen anderen Veröffentlichungswegen und -Produkten der amtlichen Statistik – keine zuordenbaren Einzelangaben veröffentlicht werden.

#### Mögliche Anonymisierungsmaßnahmen:

- Löschen oder Vergrößern von Merkmalen
- Stichprobenziehungen
- Datenveränderungen

#### Anonymisierungsgrade:

- absolut anonym
- faktisch anonym
- formal anonym

und Anonymisierung“). Im Rahmen dieser Offenlegungskontrolle werden datenschutzkritische Fälle identifiziert und mit geeigneten Verfahren aus dem Bereich der statistischen Geheimhaltung – beispielsweise der Zellspernung – bearbeitet, sodass den gesetzlichen Vorgaben entsprochen wird und die erstellten Ausgaben an die Forscherinnen und Forscher weitergegeben werden können. Diese Maßnahmen gewährleisten, dass nur absolut anonyme Ergebnisse veröffentlicht werden und aus diesen keine Rückschlüsse auf Einzelangaben gezogen werden können. Zusätzlich werden die Datennutzerinnen und -nutzer selbst durch rechtliche Vorschriften – wie die Verpflichtung auf Einhaltung der statistischen Geheimhaltung, deren Missachtung zu strafrechtlichen Konsequenzen führen kann – in die Pflicht genommen.

### Nutzungsbedingungen und Antragstellung

Grundsätzlich nutzungsberechtigt sind alle Einrichtungen mit dem Ziel unabhängiger wissenschaftlicher Forschung. Dementsprechend zählen zu diesem Kreis Hochschulen und Universitäten, aber auch Forschungsinstitute. Die Berechtigung wird bei erstmaliger Antragsstellung der Einrichtung geprüft.

Angehörige der durch die FDZ anerkannten Einrichtungen können über die FDZ-Homepage einen Online-Antrag zur Datennutzung stellen. Nach einer fachlichen Prüfung im FDZ und einem Feedbackprozess, in dessen Rahmen alle Statistischen Ämter Rückmeldungen zum Projektantrag und den beantragten Daten geben können, wird ein Nutzungsvertrag mit der jeweiligen Einrichtung geschlossen.

Die Mikrodatennutzung im FDZ der Statistischen Ämter des Bundes und der Länder ist grundsätzlich kostenpflichtig. So wird aktuell jedes beantragte Berichtsjahr einer Statistik aus dem FDZ-Angebot pro genutztem Nutzungsweg mit 250 Euro in Rechnung gestellt. Hierfür ist der Antragssteller im Gegenzug berechtigt, die entsprechende Datenbasis für eine Dauer von bis zu drei Jahren zu nutzen und in dieser Zeit entsprechend den Support sowie in Abhängigkeit vom gewählten Nutzungsweg die Infrastruktur des FDZ in Anspruch zu nehmen. Sollten die drei Jahre nicht ausreichen, so ist eine einmalige, kostenpflichtige Verlängerung der Projektlaufzeit um maximal drei weitere Jahre möglich. Für Sonderfälle, wie Peer-Review-Phasen, gibt es darüber hinaus weitere Regelungen, um den Anforderungen aus dem Wissenschaftssystem entgegenzukommen. Um Nachwuchswissenschaftlerinnen und -wissenschaftler zu fördern, existiert zudem ein Rabattsystem für Studien- beziehungsweise Abschlussarbeiten wie Masterarbeiten und Promotionen.

### Der Standort Bayern des FDZ

Hervorzuheben ist die inhaltliche und räumliche Nähe der Forschungsdatenzentren der amtlichen Statistik zur Wissenschaft. Aufgrund der regio-



### Kontaktdaten des bayerischen FDZ-Standorts

Servicenummer: +49 911 98208 6577

E-Mail: [fdz@statistik.bayern.de](mailto:fdz@statistik.bayern.de)

nenalen Infrastruktur können in jedem Standort alle Statistiken genutzt werden. Abbildung 2 enthält eine aktuelle Übersicht der regionalen Standorte.

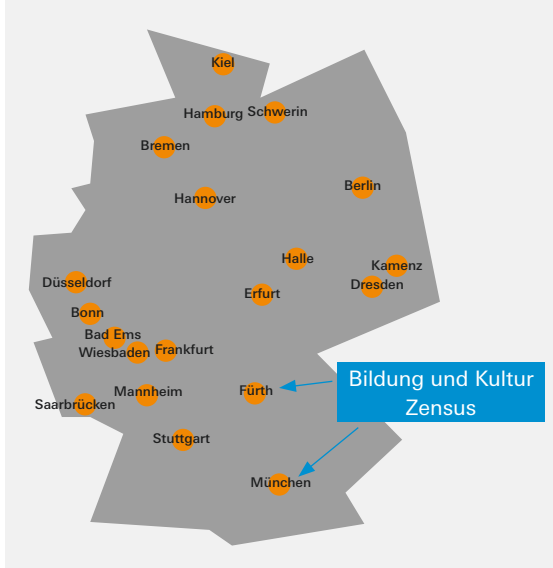
Wie im Abschnitt „Fachlich zentralisierte Datenhaltung“ beschrieben, sind die Standorte jeweils für bestimmte Statistiken fachlich zuständig. Der FDZ-Standort Bayern übernimmt die Aufbereitung und Bereitstellung der Daten aus dem Bereich Bildung und Kultur sowie des Zensus. Konkret ist Bayern aktuell für die Studierenden- und Prüfungsstatistik, die Personal- und Stellenstatistik sowie den Zensus 2011 fachlich zuständig. Die Kontaktdaten des bayerischen Standorts enthält der gleichnamige Infokasten.

Gastwissenschaftsarbetsplätze des bayerischen FDZ-Standorts sind in der Dienststelle in Fürth, Nürnberger Straße 95, und in der FDZ-Außenstelle in München, Sankt Martin-Straße 47, vorhanden. Zusätzlich wurde im Jahr 2013 eine Außenstelle mit Gastwissenschaftsarbetsplatz in München im ifo Institut, Poschingerstraße 5, in Kooperation mit dem dortigen European Business Data Center (EBDC) eingerichtet, wodurch eine noch größere Forschungsnähe erreicht werden konnte. Genutzt werden diese Arbeitsplätze überwiegend von Studierenden sowie Forscherinnen und Forschern wissenschaftlicher Einrichtungen aus Bayern. Zu den Nutzerinnen und Nutzern gehören allerdings auch Personen aus dem nahegelegenen Ausland, wie Österreich oder der Schweiz, und teilweise sogar aus dem nicht-deutschsprachigen Ausland sowie Übersee.

### Aktuelle Entwicklungen der deutschen Dateninfrastruktur

Die Entwicklung des FDZ der Statistischen Ämter der Länder und damit des FDZ-Standorts Bayern ist untrennbar verbunden mit der allgemeinen

Abb. 2  
Standorte der Forschungsdatenzentren



Fortentwicklung der Forschungsdateninfrastruktur in Deutschland:

Um Datenbestände für die gesamte Wissenschaft zu erschließen, zu vernetzen und nachhaltig nutzbar machen zu können, wurde 2020 der Verein Nationale Forschungsdateninfrastruktur (NFDI) e.V. ins Leben gerufen. Der NFDI als ein von Bund und Ländern gegründeter und aktuell geförderter Verein, hat sich zum Ziel gesetzt, ein dauerhafter Wissensspeicher für neue Forschungsfragen und Innovationen zu sein. Dabei sollen die Daten nach den FAIR-Prinzipien (Findable, Accessible, Interoperable und Reusable – also auffindbar, zugänglich, interoperabel und wiederverwendbar – zur Verfügung gestellt werden. Um eine breite Abdeckung der Wissenschaftsdisziplinen innerhalb der NFDI zu gewährleisten, werden Konsortien gefördert. In jedem Fachkonsortium engagieren sich verschiedene Einrichtungen, wie Universitäten oder Forschungszentren, und arbeiten zusammen an fachinternen sowie fachübergreifenden Themen<sup>6</sup>. Ein Konsortium, welches in diesem Beitrag Erwähnung findet, ist das Konsortium für die Sozial-, Verhaltens-, Bildungs- und Wirtschaftsdaten (KonsortSWD). Dessen Ziel ist es, die bestehende Forschungsdateninfrastruktur weiter auszubauen und gesellschaftlich relevante und sensible Daten verfügbar zu machen. Dabei baut das KonsortSWD auf bewährten wissenschaftsnahen Strukturen, insbesondere den FDZ, auf.<sup>7</sup>

Nachdem die FDZ der Statistischen Ämter des Bundes und der Länder als Vorreiter bei der erleichterten Bereitstellung forschungsrelevanter Daten mit vorangegangen waren, sind über die letzten beiden Jahrzehnte hinweg immer mehr wissenschaftliche Institutionen und Behörden diesem Beispiel gefolgt und haben ihrerseits mit der Einrichtung eigener Forschungsdatenzentren ihren Teil zum verbesserten Datenzugang zu quantitativen – und mittlerweile auch qualitativen – Forschungsdaten beigetragen. So zählt der RatSWD 37 akkreditierte FDZ aus den verschiedensten Bereichen der Sozial- und Wirtschaftsforschung. Hinzu kommen derzeit vier weitere FDZ, die sich noch in einer Phase der vorläufigen Akkreditierung befinden<sup>8</sup>. Die Interessenvertretung der FDZ

gegenüber dem RatSWD erfolgt im Rahmen des Ausschusses FDI, in dem alle vorläufig oder dauerhaft akkreditierten FDZ vertreten sind.

Die Bedeutung von Forschungsdaten für die Wissenschaft wurde auch im aktuellen Koalitionsvertrag der Parteien SPD, Bündnis 90/Die Grünen und FDP hervorgehoben. Dieser beinhaltet die Einführung eines Forschungsdatengesetzes, welches den Zugang zu Daten für Forschungszwecke verbessern soll. Der RatSWD formulierte in einem Positionspapier Eckpunkte, die bei der Ausgestaltung des Gesetzes berücksichtigt werden sollen<sup>9</sup>. Die zentrale Forderung dabei ist es, den Zugang zu Daten für die Wissenschaft zu verbessern, wobei allen Forscherinnen und Forschern die gleichen Zugangsrechte zu gewähren seien und gleichzeitig die Vertraulichkeit von Forschungsdaten gesichert werden soll. Dabei plädiert der RatSWD unter anderem für einen Datenzugang über ein Remote-Access-System, für Regelungen von erweiterten Verknüpfungsmöglichkeiten und für die Stärkung der Aufgaben der Forschungsdatenzentren. Da die Einrichtung von FDZ in den meisten Behörden eine freiwillige Aufgabe ist, soll nach Ansicht des RatSWD geprüft werden, inwieweit das Betreiben der FDZ als gesetzliche Aufgabe verankert werden kann. Dies gilt auch für die FDZ der Statistischen Ämter des Bundes und der Länder, deren Betrieb aus Sicht des Rates als eine obligatorische Kernaufgabe der Behörden verankert werden soll.

### Zukünftige Herausforderungen

Das Forschungsdatenzentrum hat sich in den letzten 20 Jahren als ein wichtiges Bindeglied zwischen der amtlichen Statistik und der Wissenschaft etabliert. Als infrastruktureller Dienstleister orientiert sich das FDZ an den Bedürfnissen der Forscherinnen und Forscher, muss dabei jedoch auch stets die rechtlichen und technischen Möglichkeiten berücksichtigen. Neben der kontinuierlichen Weiterentwicklung der Datenbestände um neue Erhebungsjahre und neue Statistiken, konnten bereits wichtige technische Neuerungen umgesetzt werden. So wurde Anfang 2022 das Online-GWAP-Buchungssystem eingeführt, über das FDZ-Nutzerinnen und -Nutzer selbstständig,

6 Nationale Forschungsdateninfrastruktur (NFDI) e.V.: Über den Verein: [www.nfdi.de/verein/](http://www.nfdi.de/verein/), abgerufen am 29.09.2022.

7 Nationale Forschungsdateninfrastruktur (NFDI) e.V.: KonsortSWD: [www.nfdi.de/konsortswd-2/](http://www.nfdi.de/konsortswd-2/), abgerufen am 29.09.2022

8 Übersicht unter [www.konsortswd.de/datenzentren/alle-datenzentren/](http://www.konsortswd.de/datenzentren/alle-datenzentren/), abgerufen am 19.10.2022

9 RatSWD (2022): Positionspapier des RatSWD: Eckpunkte für ein Forschungsdatengesetz: [www.konsortswd.de/wp-content/uploads/RatSWD-Positionspapier-Eckpunkte-fuer-ein-Forschungsdatengesetz.pdf](http://www.konsortswd.de/wp-content/uploads/RatSWD-Positionspapier-Eckpunkte-fuer-ein-Forschungsdatengesetz.pdf), abgerufen am 29.09.2022

schnell und flexibel Reservierungen von Terminen an den Gastwissenschaftsarbeitsplätzen vornehmen können. Auch ein Downloadportal für CFs und PUFs wurde zur Verfügung gestellt. Nutzerinnen und Nutzer können darüber kosten- und antragsfrei CFs und PUFs inklusive der zugehörigen Metadaten downloaden und müssen hierfür keinen gesonderten Beantragungsprozess mehr durchlaufen.

Ein Downloadportal für die SUFs befindet sich im Aufbau. Während momentan den Nutzerinnen und Nutzern die SUFs noch auf einer DVD postalisch zugesandt werden, lassen sich künftig die Daten flexibel über das Downloadportal herunterladen. Das Bayerische Landesamt für Statistik wurde mit der Umsetzung dieser Downloadlösung betraut und wird die SUFs mithilfe der in Bayern betriebenen „On-Premise-Lösung“ SecureBox bereitstellen. Mit dieser Modernisierungsmaßnahme soll den FDZ-Nutzerinnen und -Nutzern ein erhöhtes Maß an Flexibilität und Nutzerfreundlichkeit geboten werden, der Versand physischer Datenträger entfällt im Gegenzug.

Um den Forscherinnen und Forschern die Antragstellung zu erleichtern, ist ein E-Learning-Tool in Vorbereitung. In mehreren Modulen können sich interessierte Wissenschaftlerinnen und Wissenschaftler dann über das FDZ allgemein, die unterschiedlichen Zugangswege oder die Geheimhaltungs- sowie Codierungsregeln informieren. Das Tool soll dabei nicht den persönlichen Kontakt zwischen Mitarbeiterinnen und Mitarbeitern im FDZ und Forschern ersetzen, sondern als eine zusätzliche Unterstützung im Beratungsprozess dienen.

Da von Seiten der Wissenschaft, unter anderem auch vom RatSWD, schon lange ein Fernzugriff auf das Datenmaterial gefordert wird, laufen aktuell die Vorarbeiten für die Einführung eines Remote-Access-Systems. Nachdem zunächst die erforderlichen Rahmenbedingungen erarbeitet werden mussten, befindet sich das System nun in der Pilotierungsphase. Im Rahmen der aktuell geltenden rechtlichen Bestimmungen des BSatG wird zunächst der Remote-Zugriff mit Daten getestet, die momentan als SUF zur Verfügung stehen.

Nach erfolgreicher Evaluation dieses Prototyps könnte in einem Folgeschritt auch der Zugriff auf weitere anonymisierte Daten über eine gesicherte IT-Umgebung ermöglicht werden. Forscherinnen und Forscher könnten somit aus ihrer wissenschaftlichen Einrichtung heraus auf die Daten zugreifen und Analysen zeitlich flexibel durchführen, ohne an einem Gastwissenschaftsarbeitsplatz arbeiten zu müssen. Die Daten verbleiben jedoch zu jedem Zeitpunkt innerhalb der FDZ-Infrastruktur. Der Remote-Access stellt hiermit einen wichtigen Schritt zur Sicherstellung der wissenschaftlichen Wettbewerbsfähigkeit in Deutschland dar, steht dabei jedoch – wie jedes Angebot der FDZ – vor der Herausforderung, im Spannungsfeld zwischen Datenschutz- und Statistikrecht auf der einen und der Ermöglichung erweiterter Nutzungs- und Analysemöglichkeiten auf der anderen Seite die richtige Lösung zu finden.

Als eine der wichtigsten Datenquellen für politische, gesellschaftliche oder wirtschaftliche Entscheidungen gilt der Zensus. Bereits die FDZ-Daten des Zensus 2011 wurden von Forscherinnen und Forschern für ihre Analysen verwendet. Zukünftig sollen auch die Daten des Zensus 2022 für die Wissenschaft in Form von FDZ-Daten aufbereitet und zur Verfügung gestellt werden. Als fachlich zuständiger FDZ-Standort ist Bayern hierbei federführend. Aktuell laufen in einer hierfür gemeinsam mit anderen FDZ-Standorten sowie der Zensus-Fachseite ins Leben gerufenen Projektgruppe die Vorarbeiten, um die Daten auch der Wissenschaft zugänglich machen zu können. Neben der eigentlichen Datenaufbereitung für die verschiedenen Statistikprogramme spielt hier auch die Erstellung von Metadaten und die Bereitstellung von Begleitdokumenten zur Dokumentation der komplexen Datenstrukturen eine wichtige Rolle.

Neben technischen Gegebenheiten und den Datenprodukten wandeln sich auch die zum Einsatz kommenden Geheimhaltungs- und Analyseverfahren. Zukünftig werden Verfahren, welche zur Sicherstellung der statistischen Geheimhaltung eine stochastische Datenveränderung und keine Datensperre vorsehen (Cell-Key-Methode),



voraussichtlich eine zunehmend größere Rolle spielen. Damit einher geht die Hoffnung, den Geheimhaltungsprozess (teil-)automatisieren zu können, das Recht auf informationelle Selbstbestimmung abzusichern und zugleich den Datenutzerinnen und -nutzern ein Mehr an Informationen gegenüber dem Status Quo verfügbar machen zu können.

### **Zusammenfassung und Ausblick**

Die Forschungsinfrastruktur in Deutschland hat sich innerhalb der letzten beiden Jahrzehnte massiv gewandelt und einen Prozess hin zu einer erleichterten Verfügbarkeit zahlreicher Datenbestände von Wissenschaft und Verwaltung aus den unterschiedlichsten Themenbereichen geführt. Dies gilt ebenso für die amtliche Statistik in Deutschland, die mit dem FDZ des Statistischen Bundesamtes und dem FDZ der Statistischen Ämter der Länder – welches in diesem Jahr sein 20. Gründungsjubiläum feiert – eine Vorreiterrolle eingenommen hat. Wichtige Etappen, die seit damals erfolgreich absolviert wurden, waren die dauerhafte Etablierung des FDZ seitens der Statistischen Ämter der Länder, die Gründung einer Reihe von FDZ-Außenstellen innerhalb wissenschaftlicher Einrichtungen sowie der nutzerorientierte Ausbau des Daten- und Dienstleistungsangebots. Auch der FDZ-Standort des Bayerischen Landesamts für Statistik hat sich in dieser Zeit zu

einem wichtigen Standort für die Wissenschaft entwickelt: Neben einer Ausweitung des Datenangebots ist hier unter anderem die Gründung der Außenstelle im renommierten Münchner ifo Institut zu nennen. Trotz der Verlagerung des Landesamts von München nach Fürth, werden weiterhin die Gastwissenschaftsarbeitsplätze an beiden Standorten in München für FDZ-Nutzerinnen und -Nutzer angeboten. Die Übernahme der fachlichen und organisatorischen Zuständigkeit für die komplexen Daten des Zensus 2011 und 2022 ist ebenfalls prägend für den FDZ-Standort Bayern als Datenschnittstelle zwischen amtlicher Statistik und empirisch forschender Wissenschaft.

Doch auch die kommenden Jahre werden im Bereich der Forschungsdateninfrastruktur im Ganzen, genau wie für den FDZ-Standort im Bayerischen Landesamt für Statistik, weitere Herausforderungen mit sich bringen: Die sich stetig wandelnden Bedürfnisse der Wissenschaft und die immer besser werdenden technischen Möglichkeiten erfordern eine kontinuierliche Anpassung und Weiterentwicklung in den FDZ. So wird zum einen von Bedeutung sein, das bisher Erreichte weiterhin zu sichern und die vorhandenen Stärken auszubauen, und zum anderen die Modernisierung der amtlichen Dateninfrastruktur weiter voranzutreiben.