

10 Jahre „Wege zu einer besseren informationellen Infrastruktur“

Die Statistischen Ämter des Bundes und der Länder existieren seit mehr als 6 Jahrzehnten und sind die bedeutendsten Datenproduzenten in Deutschland. So breit gefächert die Bereiche sind, in denen hier Statistiken erhoben werden, so verschieden sind auch die Nutzer der amtlichen Daten. Zu den Hauptanwendern der Daten zählen die Politik, die Medien und Bürger, nicht zuletzt aber auch die Wissenschaft. Aus den immer komplexer werdenden sozialen, ökonomischen und umweltrelevanten Herausforderungen in unserer Gesellschaft ergeben sich für die Wissenschaft immer differenziertere Forschungsfragen. Aggregierte Ergebnisse oder Tabellen reichen zur Beantwortung solcher Fragestellungen häufig nicht mehr aus. Von der Wissenschaft werden Einzeldaten benötigt. Der komfortable Zugang der Wissenschaft zu faktisch anonymen Einzeldaten über die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder ist mittlerweile – sowohl für die Forschung als auch für die amtliche Statistik selbst – zu einer Selbstverständlichkeit geworden. Aber was heute selbstverständlich ist, erschien noch im Jahr 2001 kaum vorstellbar. Die Nachfrage nach Einzeldaten für Forschungszwecke war auch vor 10 Jahren nicht neu, die Hürden, die es in der Vergangenheit von der Wissenschaft beim Zugang zu diesen Informationen zu nehmen galt, waren allerdings hoch. Vor 10 Jahren – im März 2001 – wurde ein Gutachten vorgelegt, das auf diesen Mangel bei der „informationellen Infrastruktur“ aufmerksam machte. Dieses Gutachten stellt einen bedeutenden Meilenstein für die Einrichtung von Forschungsdatenzentren (FDZ) bei Datenproduzenten wie den statistischen Ämtern dar.

Der folgende Beitrag nimmt das zehnjährige Jubiläum des Gutachtens zum Anlass, die Entwicklung der informationellen Infrastruktur in Deutschland nachzuzeichnen und richtet dabei den Blick insbesondere auf die amtliche Statistik.

Hintergrund

Am 13. März 2001 wurde ein im Auftrag des Bundesministeriums für Bildung und Forschung (BMBF) erstelltes Gutachten vorgelegt, das sich mit den damaligen Beschränkungen für die empirisch forschende Wissenschaft beim Zugang zu statistischen Einzeldaten beschäftigte. Neben einer bloßen Bestandsaufnahme der damaligen Situation enthielt das Gutachten auch Lösungsvorschläge, die in der Folge von Datenproduzenten wie den Statistischen Ämtern des Bundes und der Länder aufgegriffen wurden. Ein weiterer zentraler Punkt des Gutachtens war die Empfehlung einer generellen Verbesserung von Kommunikation und Kooperation zwischen Wissenschaft und Statistik. Das Gutachten mit dem Titel „Wege zu einer besseren informationellen Infrastruktur“ war das Ergebnis der 1999 vom BMBF eingesetzten „Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik“ (KVI). Im Auftrag der damaligen Bundesministerin für Bildung und Forschung Bulmahn sollte die KVI mit dem Gutachten klären, ob die informationelle Infrastruktur in Deutschland noch zeitgemäß ist und den gewandelten Anforderungen der Wissenschaft gerecht wird¹⁾.

¹⁾ Vgl. Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (Hrsg.): Wege zu einer besseren informationellen Infrastruktur, Baden-Baden, 2001, S. 9 f.

Anforderungen der Wissenschaft

Heute stellt sich in diesem Zusammenhang fast automatisch die Frage, inwiefern sich die Anforderungen der Wissenschaft an die amtliche Statistik in der Vergangenheit gewandelt haben und worin Gründe für diesen Wandel lagen.

Das Bundesstatistikgesetz und die jeweiligen Landesstatistikgesetze sehen eine Verpflichtung der amtlichen Statistik zur Veröffentlichung und Darstellung von statistischen Ergebnissen vor. Daher fasst die amtliche Statistik die statistischen Informationen als öffentliches Gut auf und stellt diese der breiten Öffentlichkeit über das Internet (z. B. als Online-Standardprodukt) kostenfrei zur Verfügung. Daneben werden von den statistischen Ämtern für spezielle Zielgruppen auch Standardprodukte in gedruckter Form oder auf Datenträgern, in denen statistische Daten in besonderer Weise zusammengetragen und vertiefend dargestellt bzw. analysiert werden, erstellt. Diese werden zur Deckung der Herstellungs- und Versandkosten in der Regel gegen Entgelt abgegeben. In beiden Fällen erhält der Nutzer die Ergebnisse in aggregierter und standardisierter Form. Sofern diese Standardangebote nicht ausreichen, können bei den statistischen Ämtern Sonderauswertungen in Auftrag gegeben werden. Solche Sonderauswertungen sind allerdings aufwändig und damit für den Nutzer auch teuer. Zudem reichen die erzeugten Ergebnisse zur Beant-

wortung komplexer Forschungsfragen oftmals nicht aus. Für eine tiefgehende Beurteilung gesellschaftlicher und ökonomischer Zusammenhänge bedarf es meist des Blickes „hinter“ einen aggregierten Wert. Dies gilt z. B. für die Beantwortung von Fragestellungen, die sich mit bestimmten Teilgruppen aus Wirtschaft und Gesellschaft beschäftigen²⁾. Darüber hinaus gewinnen aber auch Daten für kleinräumige Untersuchungen oder georeferenzierte Daten immer mehr an Bedeutung.

Um aus dem vorhandenen Datenmaterial die benötigten Informationen zu erhalten, bedarf es einer Vielzahl unterschiedlicher wissenschaftlicher Methoden³⁾. Für die meisten heute benutzten statistischen Methoden und Verfahren ist die Verwendung aggregierter Ergebnisse, die in Form von Tabellen oder Makrodaten weitergegeben werden, aber schlichtweg ungeeignet. Zur umfassenden Beantwortung aktueller Fragestellungen werden von der Wissenschaft Einzeldaten benötigt.

Technische Rahmenbedingungen für die Nutzung von Einzeldaten

Die benötigten Einzeldaten sind in großem Umfang bei Datenproduzenten wie den Statistischen Ämtern des Bundes und der Länder, der Bundesagentur für Arbeit oder den Rentenversicherungsträgern vorhanden. Speziell die Einzeldaten der amtlichen Statistik umfassen dabei eine Vielzahl verschiedener Themengebiete. So werden neben Wirtschafts-, Bevölkerungs- und Sozialstatistiken beispielsweise auch Informationen über energie- und umweltpolitische Sachverhalte oder Daten zur Landwirtschaft, Bildung und Rechtspflege erhoben.

Die Einzeldaten der großen Datenproduzenten haben erhebliche Vorteile gegenüber Erhebungen, wie sie etwa von Forschungsinstituten durchgeführt werden. So lassen sich dank regelmäßig wiederkehrender Erfassung und oftmals über lange Zeiträume konstant bleibende Merkmalskataloge Veränderungen im Zeitablauf besonders gut abbilden. Die Daten der amtlichen Statistik sind in vielen Fällen Sekundärerhebungen. Das heißt, dass die Informationen nicht von den statistischen Ämtern selbst erfasst werden, sondern vielfach aus Verwaltungsprozessen anderer Institutionen stammen. So erfolgt in

Deutschland beispielsweise die Meldung von Eheschließungen, Geburten oder Sterbefällen bei den Standesämtern. Die dort erfassten Informationen, die für die Erstellung der Statistiken zur natürlichen Bevölkerungsbewegung von Bedeutung sind, werden in der Folge von den Standesämtern an die statistischen Landesämter weitergeleitet⁴⁾. Ebenso werden z. B. auch Angaben für die Erstellung von Steuerstatistiken nicht durch Direktbefragungen der statistischen Ämter gewonnen. Die für die Durchführung der Steuerstatistiken benötigten Angaben fallen vielmehr bei den zuständigen Finanzämtern im Rahmen der Besteuerung an und werden von diesen an die statistischen Landesämter übermittelt. Antwortausfälle, wie es sie beispielsweise bei vielen sozialwissenschaftlichen Erhebungen gibt, sind hier unbekannt⁵⁾.

Viele amtliche Statistiken stellen Vollerhebungen dar oder basieren – wie beispielsweise der Mikrozensus – auf sehr großen Stichproben. Im Ergebnis umfassen die amtlichen Daten eine hohe Anzahl an Fällen. Kurz gesagt, Daten der amtlichen Statistik sind Massendaten. Waren vor ein paar Jahrzehnten Datenbestände, die zum Teil aus mehreren Millionen Einzeldatensätzen bestehen, nur mit großem Aufwand handhabbar und mussten umständlich am Großrechner ausgewertet werden, spielen solche Probleme im Zeitverlauf eine immer geringere Rolle. Im Jahr 2001, also dem Jahr, in dem das Gutachten vorgelegt wurde, stellte auch ein großes Datenvolumen für die Auswertung von Einzeldaten keine wirkliche Hürde mehr dar. Auch umfangreiche Einzeldatensätze konnten nun relativ einfach am PC analysiert werden. Hierzu hatte, neben einer verbesserten Hardware, auch die Entwicklung und Verbesserung entsprechender Softwarepakete ihren Beitrag geleistet. Seit den Achtzigerjahren des letzten Jahrhunderts sind Programmpakete zur Analyse großer Datenmengen nicht mehr nur für Großrechner, sondern auch für den PC verfügbar. Die statistischen Standardprogramme verfügen neben Kommandozeilen zur Eingabe einer Programmsyntax heute in der Regel auch über intuitive Menüsteuerungen. Eine Auswertung von Einzeldatensätzen mit diesen Statistikpaketen stellt somit selbst für Nutzer, die im Umgang mit den programmeigenen Syntax-Codes nicht vertraut sind, kein großes Problem dar. Zugleich gibt es mittlerweile frei verfügbare und kostenlose Alternativen zu den kommerziellen Statistikpaketen. Mit den gewonnenen technischen Möglichkeiten für die Auswertung amtlicher Mikrodaten stieg im Ergebnis auch die Nachfrage der Wissenschaft nach diesen Daten.

Rechtliche Rahmenbedingungen für die Bereitstellung von Einzeldaten

Neben dem Bedarf der Wissenschaft und den vorhandenen technischen Möglichkeiten ist im Zusammenhang mit dem Zugang zu Einzeldaten der amtlichen Statistik

2) Vgl. Zühlke, S.; Zwick, M.; Scharnhorst, S.; Wende, T.: Die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder, In: Statistische Analysen und Studien NRW, Band 13, Düsseldorf, 2004, S. 3.

3) Vgl. Bauer, P.: Das Forschungsdatenzentrum der Statistischen Landesämter – Ein Beitrag zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik. In: Rolf, G.; Zwick, M.; Wagner, G. (Hrsg.): Fortschritte der informationellen Infrastruktur in Deutschland – Festschrift für Johann Hahlen zum 65. Geburtstag und Hans-Jürgen Krupp zum 75. Geburtstag, Baden-Baden, 2008, S. 159.

4) Vgl. Richter, A.: Einzeldaten der Bevölkerungsstatistik – Die natürliche Bevölkerungsbewegung in Deutschland. In: Statistische Monatshefte Rheinland-Pfalz, Heft 7/2006, S. 410 ff.

5) Vgl. Wirth H.; Müller, W.: Mikrodaten der amtlichen Statistik – Ihr Potenzial in der empirischen Sozialforschung. In: Diekmann, A. (Hrsg.): Methoden der Sozialforschung, Wiesbaden, 2006, S. 98.

Stationen auf dem Weg zu den Forschungsdatenzentren (FDZ) der amtlichen Statistik



HSL 15/11

noch ein weiterer, sehr wesentlicher Aspekt zu berücksichtigen. Die entsprechenden rechtlichen Rahmenbedingungen müssen gegeben sein. Maßgeblich für die Übermittlung amtlicher Daten ist neben dem Datenschutzgesetz das Bundesstatistikgesetz (BStatG)⁶⁾. Mit der ersten Fassung des BStatG von 1953 ergaben sich für die Weitergabe amtlicher Daten an die Wissenschaft kaum rechtliche Einschränkungen. Gegen einen verbreit-

6) Gesetz über die Statistik für Bundeszwecke.

7) Vgl. Zühlke, S.; Zwick, M.; Scharnhorst, S.; Wende, T.: Die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder, In: Statistische Analysen und Studien NRW, Band 13, Düsseldorf, 2004, S. 3.

teten Datenzugang sprachen aber bis in die Siebzigerjahre die technischen Beschränkungen, sodass die Zahl der Datennutzungen zu dieser Zeit überschaubar blieb⁷⁾. Während sich die technischen Möglichkeiten ganz langsam verbesserten, kam es im Jahr 1980 zu einer Novellierung des BStatG mit der Folge, dass von den statistischen Ämtern nur noch vollständig anonyme Daten in Form von aggregierten Werten oder Tabellen an die Wissenschaft weitergegeben werden durften. Auf dem Weg zu einer besseren informationellen Infrastruktur ging man damit einen großen Schritt zurück.

In der Bundesrepublik gelten sowohl das Recht auf informationelle Selbstbestimmung als auch die Freiheit von Wissenschaft und Forschung⁸⁾. Diesem scheinbaren Widerspruch wurde mit der erneuten Novellierung des Bundesstatistikgesetzes im Jahr 1987 Rechnung getragen. § 16 Abs. 6 erlaubte es den statistischen Ämtern von Bund und Ländern nun, der unabhängigen wissenschaftlichen Forschung den Zugang zu Einzeldaten für Forschungsprojekte zu ermöglichen, sofern die Daten „nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft“ zugeordnet werden können – also faktisch anonym sind – und die Empfänger entsprechend verpflichtet wurden. Seit 1987 gab es damit eine rechtliche Grundlage für den Zugang der Wissenschaft zu Einzeldaten der amtlichen Statistik, das sogenannte „Wissenschaftsprivileg“.

Zugangsmöglichkeiten in den Neunzigerjahren

Auch wenn die damaligen technischen und rechtlichen Rahmenbedingungen für einen verbreiteten Zugang zu amtlichen Einzeldaten für Forschungszwecke sprachen, blieben die Möglichkeiten für die Wissenschaft in der Praxis weiter eingeschränkt. Von der amtlichen Statistik wurden für Forschungsprojekte zwar anonymisierte Einzeldaten zur Verfügung gestellt, allerdings war die Datenbereitstellung teilweise mit hohen Kosten verbunden und blieb auf wenige Statistiken beschränkt. Zumeist waren dies Statistiken, die im Zusammenhang mit Haushalts- oder personenbezogenen Erhebungen standen, wie beispielsweise der Mikrozensus oder die Einkommens- und Verbrauchsstichprobe. Die Bereitstellung erfolgte in der Regel in Form faktisch anonymer Scientific-Use-Files für die Off-Site-Nutzung, also die Nutzung außerhalb der Räume der amtlichen Statistik. Für Einzeldaten aus den Wirtschaftsstatistiken, bei denen sich insbesondere bei tiefer räumlicher oder sachlicher Gliederung die faktische Anonymität außerhalb der geschützten Räume der amtlichen Statistik kaum sicherstellen lässt, konnten keine Datensätze für die Off-Site-Nutzung zur Verfügung gestellt werden. In einigen Fällen bot sich für Wissenschaftler die Möglichkeit, auf Grundlage ge-

sonderter Vereinbarungen mit einem statistischen Landesamt Einzeldaten innerhalb der geschützten Räume der amtlichen Statistik zu nutzen. Die Auswertung blieb dabei jedoch auf die Daten des jeweiligen Landesamtes beschränkt und war folglich nur für ein einzelnes Bundesland durchführbar⁹⁾. Die Möglichkeit, amtliche Mikrodaten für das gesamte Bundesgebiet an einem Ort auszuwerten, bestand zu dieser Zeit nicht. Der Zugang zu Einzeldaten der amtlichen Statistik blieb somit in der Praxis mit großen Einschränkungen verbunden.

Nicht nur die Beantwortung gesellschaftlich bedeutender Fragestellungen, sondern auch die internationale Wettbewerbsfähigkeit der empirischen Wirtschaftsforschung in Deutschland standen immer mehr auf dem Spiel. Zum Ende der Neunzigerjahre wurden daher insbesondere bei den Wirtschaftswissenschaftlern die Forderungen nach Lösungen für diese Beschränkungen immer lauter. Ein in dieser Zeit aufgekommener Vorschlag war es, Nutzern den Datenzugang in den statistischen Ämtern als sogenannte „One-Dollar-Men“ zu ermöglichen. Nach diesem Vorschlag hätten Datennutzer bei einem statistischen Amt für einen „symbolischen Dollar“ angestellt und entsprechend verpflichtet werden sollen. In der Folge sollten sie als „Mitarbeiter“ eines statistischen Amtes die Möglichkeit haben, Einzeldaten in den Räumen der amtlichen Statistik auszuwerten. Dieser Vorschlag wurde dann allerdings von den Datenschutzbeauftragten abgelehnt¹⁰⁾. Und mehr als 10 Jahre später erscheint diese Idee, auch nicht zuletzt mit Blick auf die allgegenwärtige Frage nach einer stabilen Finanzierung der heutigen Forschungsdatenzentren, eher kurios als praxisnah.

Die Lage wird bewertet

Bei der Diskussion um die begrenzten Zugangsmöglichkeiten zu Einzeldaten wurde darauf aufmerksam gemacht, dass ohne einen entsprechenden Datenzugang eine erfolgreiche Wirtschaftsforschung in der Bundesrepublik unmöglich ist¹¹⁾. Im Jahr 1998 setzte sich unter anderem der Wissenschaftsrat in einer Stellungnahme zu den Wirtschaftsforschungsinstituten mit diesem Problem auseinander¹²⁾.

Was ist der Wissenschaftsrat und welche Aufgaben hat er?

Der **Wissenschaftsrat (WR)** wurde 1957 von Bund und Ländern eingerichtet, um einen Überblick über die wissenschaftliche Arbeit in Deutschland zu geben und die Bundes- bzw. Landesregierungen bei der inhaltlichen und strukturellen Entwicklung von Hochschulen, Wissenschaft und Forschung zu beraten. Der WR setzt sich aus Wissenschaftlern, Persönlichkeiten des öffentlichen Lebens sowie Vertretern der Bundes- und Landesregierungen zusammen. (<http://www.wissenschaftsrat.de>)

8) Vgl. ebd. S. 4.

9) Vgl. Wagner, J.: FiDASt Reloaded. In: Rolf, G.; Zwick, M.; Wagner, G. (Hrsg.): Fortschritte der informationellen Infrastruktur in Deutschland – Festschrift für Johann Hahlen zum 65. Geburtstag und Hans-Jürgen Krupp zum 75. Geburtstag, Baden-Baden, 2008, S. 370 f.

10) Vgl. Bauer, P.: Das Forschungsdatenzentrum der Statistischen Landesämter – Ein Beitrag zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik. In: Rolf, G.; Zwick, M.; Wagner, G. (Hrsg.): Fortschritte der informationellen Infrastruktur in Deutschland – Festschrift für Johann Hahlen zum 65. Geburtstag und Hans-Jürgen Krupp zum 75. Geburtstag, Baden-Baden, 2008, S. 163.

11) Vgl. Hauser, R.; Wagner, G.; Zimmermann, F.: Memorandum: Erfolgsbedingungen empirischer Wirtschaftsforschung und empirisch gestützter Wirtschafts- und sozialpolitischer Beratung – IZA Discussion Paper No. 14, Bonn, 1998, S. 3.

12) Vgl.: Wissenschaftsrat (Hrsg.): Stellungnahme zu den Wirtschaftsforschungsinstituten der Blauen Liste in den alten Bundesländern – Allgemeine Gesichtspunkte, Berlin, 1998, S. 13 ff.

Im Jahr 1999 wurde vom Bundesministerium für Bildung und Forschung die **Kommission zur Verbesserung der informationellen Infrastruktur** (KVI) eingesetzt. Sie hatte die Aufgabe, eine umfassende Bestandsaufnahme der Situation durchzuführen und Vorschläge für eine Verbesserung vorzulegen. Das Ergebnis war das vor 10 Jahren vorgelegte Gutachten „Wege zu einer besseren informationellen Infrastruktur“, in dem die KVI auf die Notwendigkeiten von erleichterten Zugangswegen zu Einzeldaten der amtlichen Statistik hinwies. Die KVI machte auch darauf aufmerksam, dass das vorhandene Informationsangebot in Deutschland zwar durchaus umfassend, aber gleichzeitig auch sehr unübersichtlich sei. Die amtlichen Daten würden in erster Linie für administrative Zwecke erhoben, sachgerechte Informationen für wissenschaftliche Datennutzer würden vielfach fehlen. Neben geregelten Zugangswegen bedürfe es daher auch entsprechender Datendokumentationen, die der Wissenschaft Informationen über Erhebungs- und Aufbereitungsverfahren liefern.

Der Rat für Sozial- und Wirtschaftsdaten

Die KVI wies in ihrem Gutachten auch darauf hin, dass es einer Verbesserung bei der Kommunikation und Kooperation zwischen Wissenschaft und Statistik bedürfe. In diesem Zusammenhang empfahl die KVI die Einrichtung eines Rates für Sozial- und Wirtschaftsdaten (RatSWD), der die „Aufgabe (haben soll), die Dateninfrastruktur in Deutschland zu bewerten, weiterzuentwickeln sowie für Bund und Länder Empfehlungen zum Programm der wissenschaftstragenen Statistik und dessen Finanzierung zu erarbeiten“¹³⁾.

Nachdem im Jahr 2001 ein Gründungsausschuss für den RatSWD eingerichtet worden war, wurde der Rat im Jahr 2004 erstmalig vom Bundesministerium für Bildung und Forschung berufen. Ein wesentliches Ziel des RatSWD ist es heute, „die Situation der empirischen Sozial- und Wirtschaftswissenschaften nachhaltig zu verbessern und zu ihrer internationalen Wettbewerbsfähigkeit beizutragen“¹⁴⁾. Der RatSWD setzt sich sowohl aus Vertretern der Wissenschaft als auch der Datenproduzenten zusammen. Die Statistischen Ämter des Bundes und der Länder sind als größte Datenproduzenten in Deutschland von Beginn an Mitglieder im Rat. Das Statistische Bundesamt wird durch seinen Präsidenten Roderich Egeler vertreten. Die statistischen Ämter der Länder vertritt seit Gründung des Rates der Präsident des Hessischen Statistischen Landesamtes, Herr Eckart Hohmann.

13) Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (Hrsg.): Wege zu einer besseren informationellen Infrastruktur, Baden-Baden, 2001, S. 27.

14) <http://www.ratswd.de/rat/aufgaben.php> (Abruf vom 27. Januar 2011).

Einrichtung von Forschungsdaten- und Datenservicezentren

Eine Aufgabe des RatSWD ist die Weiterentwicklung und Bewertung der informationellen Infrastruktur in Deutschland. Hierbei kommt bereits dem Gründungsausschuss des Rates eine entscheidende Rolle zu, denn dieser hat die Einrichtung von Forschungsdatenzentren (FDZ) bei den Datenproduzenten angeregt. Zu den ersten FDZ, die in Deutschland zwischen 2001 und 2004 eingerichtet wurden, zählen neben dem FDZ des Statistischen Bundesamtes und dem der statistischen Landesämter auch das FDZ der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung sowie das FDZ der Deutschen Rentenversicherung Bund. Daneben wurden auch zwei Datenservicezentren (DSZ) eingerichtet. Eines bei der Gesellschaft Sozialwissenschaftlicher Infrastruktureinrichtungen (GESIS) und eines beim Institut zur Zukunft der Arbeit (IZA). Allen ist gemein, dass diese in ihrer Startphase vom BMBF mit Anschubfinanzierungen gefördert wurden. Mittlerweile sind die genannten FDZ und DSZ nicht mehr die einzigen Einrichtungen dieser Art in Deutschland. Heute, 10 Jahre nach Vorlage des Gutachtens der KVI, gibt es bereits 15 FDZ und 3 DSZ.

Die FDZ und DSZ werden vom RatSWD akkreditiert. Für die Akkreditierung müssen von den FDZ und DSZ bestimmte Kriterien erfüllt werden. Welche Anforderungen

Was sind Forschungsdatenzentren, was sind Datenservicezentren?

Forschungsdatenzentren (FDZ) sind Einrichtungen von Datenproduzenten. Unter Wahrung der rechtlichen Bestimmungen haben FDZ die Aufgabe, der Wissenschaft einen einfachen und transparenten Zugang zu Einzeldaten zu ermöglichen. Dies kann durch die Bereitstellung von faktisch anonymisierten Scientific-Use-Files oder die Bereitstellung faktisch anonymisierter Daten an einem Gastwissenschaftlerarbeitsplatz erfolgen. Lassen sich Daten nicht angemessen anonymisieren, besteht in der Regel die Möglichkeit, Auswertungen über eine kontrollierte Datenfernverarbeitung durchzuführen. In diesem Fall sendet der Datennutzer ein Auswertungsprogramm (Syntax) an das FDZ und erhält in der Folge vollständig anonymisierte Analyseergebnisse.

Datenservicezentren (DSZ) sind Einrichtungen außerhalb der Institutionen großer Datenproduzenten, die empirisch forschende Sozial- und Wirtschaftswissenschaftler durch bestimmte Serviceleistungen unterstützen. Zu den Serviceleistungen können beispielsweise die Erstellung umfassender Datendokumentationen, die Archivierung von Daten oder die Schulung von Nutzern im Umgang mit Daten zählen.

Forschungsdaten- und Datenservicezentren in Deutschland

Forschungsdatenzentren (FDZ)	Standorte
FDZ der statistischen Ämter der Länder	Wiesbaden, Kiel, Schwerin, Hamburg, Bremen, Berlin Mitte, Berlin Alt-Friedrichsfelde, Potsdam, Hannover, Halle (Saale), Düsseldorf, Kamenz, Dresden, Erfurt, Bad Ems, Saarbrücken, Stuttgart, München
FDZ des Statistischen Bundesamtes	Wiesbaden, Berlin, Bonn
FDZ der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung	Nürnberg ¹⁾
FDZ der Deutschen Rentenversicherung Bund	Berlin, Würzburg
FDZ im Bundesinstitut für Berufsbildung	Bonn
FDZ am Institut zur Qualitätsentwicklung im Bildungswesen	Berlin
FDZ des Sozio-oekonomischen Panels	Berlin
FDZ ALLBUS bei GESIS	Berlin, Bonn, Köln, Mannheim
FDZ „Internationale Umfrageprogramme“ bei GESIS	Berlin, Bonn, Köln, Mannheim
FDZ „Wahlen“ bei GESIS	Berlin, Bonn, Köln, Mannheim
FDZ Survey of Health, Ageing and Retirement in Europe — SHARE	Mannheim
FDZ Deutscher Alterssurvey	Berlin
FDZ PsychData des Leibniz-Zentrums für Psychologische Information und Dokumentation	Trier
FDZ Panel Analysis of Intimate Relationships and Family Dynamics — PAIRFAM	Bremen, Chemnitz, Mannheim
FDZ RWI	Essen

Datenservicezentren (DSZ)	Standorte
German Microdata Lab (GML) Servicezentrum für Mikrodaten des Leibniz-Instituts für Sozialwissenschaften (GESIS) / MISSY	Berlin, Bonn, Köln, Mannheim
Internationales Datenservicezentrum des Forschungsinstituts zur Zukunft der Arbeit (IZA)	Bonn
Datenservicezentrum Betriebs- und Organisations-Daten (DSZ-BO)	Bielefeld

Quelle: <http://www.ratswd.de/dat/fdz.php> (Abruf vom 10. Februar 2011).

¹⁾ In diesen Tagen läuft das Pilot-Projekt „PFiF“ (Projekt FDZ im FDZ) an, bei dem in einem ersten Versuch der Zugang zu Daten des FDZ IAB auch in den Standorten Berlin, Bremen, Dresden und Düsseldorf des FDZ der statistischen Ämter der Länder getestet wird. Das FDZ des IAB nutzt damit als erstes das dichte Netzwerk von Standorten des Forschungsdaten-zentrums der statistischen Ämter der Länder.

muss ein Forschungsdaten-zentrum also grundsätzlich erfüllen?

Die Hauptaufgabe eines FDZ ist ein erleichterter Zugang zu Daten für wissenschaftliche Nutzer unter Wahrung der geltenden (datenschutz-)rechtlichen Regelungen. Das FDZ stellt damit eine Serviceeinrichtung eines Datenproduzenten dar. In der Bundesrepublik gilt die Freiheit von Wissenschaft und Forschung. Somit darf durch ein FDZ keine inhaltliche Bewertung eines wissenschaftlichen Projektes vorgenommen werden. Alle Nutzer sind folglich – unabhängig von ihrer jeweiligen Fragestellung – gleichberechtigt zu behandeln.

Um die Mikrodaten, die ausgewertet werden sollen, korrekt interpretieren zu können, bedarf es Informationen, die eine korrekte Deutung von statistischen Daten ermöglichen, wie beispielsweise die Darstellung von Er-

hebungsmethodiken oder Merkmalsdefinitionen¹⁵⁾. Eine Aufgabe von FDZ ist es daher ebenfalls, solche Zusatzinformationen – sogenannte Metadaten – für die wissenschaftlichen Nutzer zu erstellen und diesen zugänglich zu machen.

Für die Akkreditierung der FDZ spielt des Weiteren die aktive Beteiligung an der wissenschaftlichen Diskussion eine wichtige Rolle. Die FDZ verpflichten sich, diese Kriterien zu erfüllen und sind dem RatSWD jährlich rechen-schaftspflichtig¹⁶⁾.

Das FDZ der statistischen Ämter der Länder

Das FDZ des Statistischen Bundesamtes und das der statistischen Ämter der Länder gehören zu den FDZ der ersten Stunde. Dabei ist das FDZ der statistischen Landesämter mit seinen bundesweit 18 Standorten¹⁷⁾ das größte FDZ in Deutschland. Das FDZ ist eine Arbeitsgemeinschaft aller 14 statistischen Landesämter. Aufgaben und Arbeitsweise sind im Rahmen einer Kooperationsvereinbarung festgelegt, die im April 2002 mit der Gründung des FDZ verabschiedet wurde.

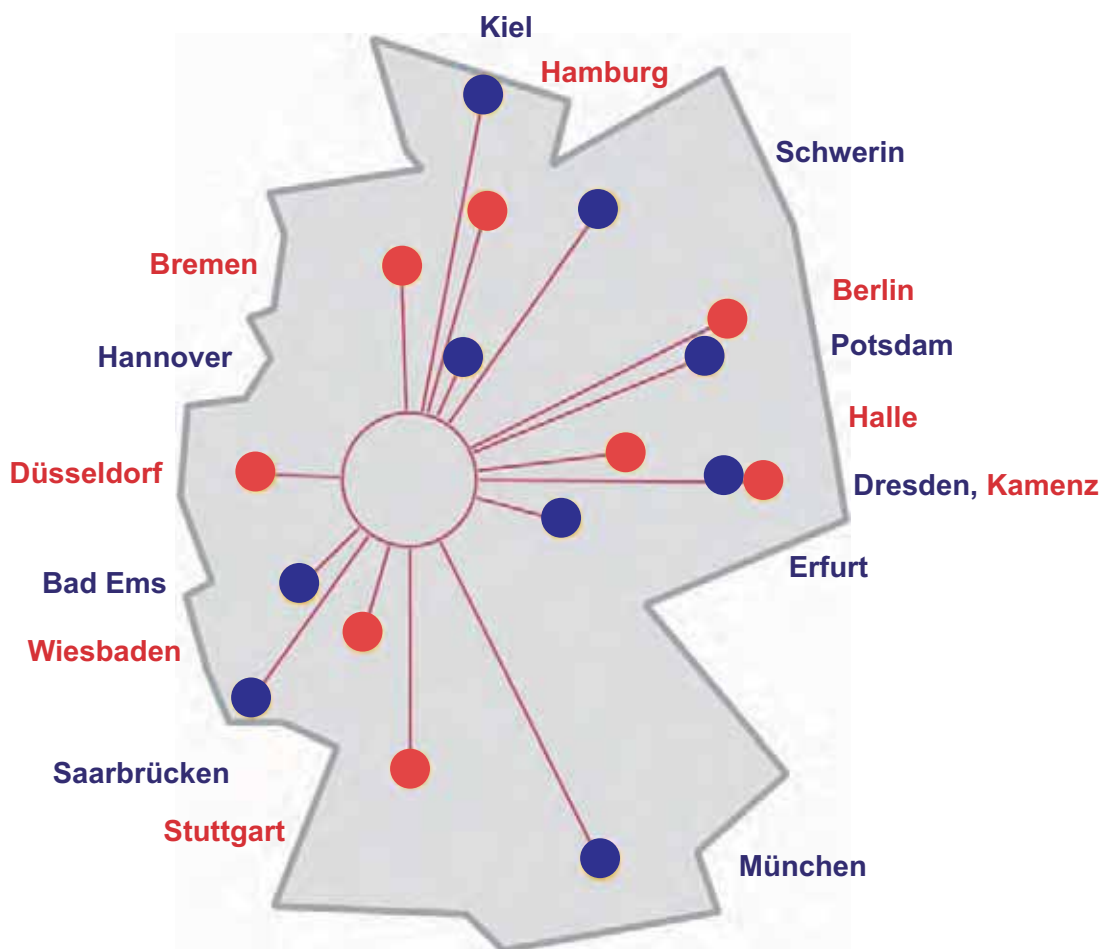
Mit der Einrichtung eines gemeinsamen FDZ haben die statistischen Ämter der Länder als größte Datenproduzen-

¹⁵⁾ Vgl. Richter, A.; Weil, S.: Metadaten – Eine Grundlage für die Auswertung amtlicher Statistiken durch die Wissenschaft. In: Statistische Monatshefte Rheinland-Pfalz, Heft 1/2005, S. 14 ff.

¹⁶⁾ Vgl. RatSWD (Hrsg.): Kriterien des Rates für Sozial- und Wirtschaftsdaten (RatSWD) für die Einrichtung der Forschungsdaten-Infrastruktur, Berlin, 2010.

¹⁷⁾ Standorte: Wiesbaden, Kiel, Schwerin, Hamburg, Bremen, Berlin Mitte, Berlin Alt-Friedrichsfelde, Potsdam, Hannover, Halle (Saale), Düsseldorf, Kamenz, Dresden, Erfurt, Bad Ems, Saarbrücken, Stuttgart, München.

Regionale Standorte des FDZ der statistischen Ämter der Länder



HSL 16/11

ten Deutschlands eine Dateninfrastruktur geschaffen, die sich nicht nur durch ein dichtes regionales Netzwerk, sondern vor allem auch durch ein fachlich breites Angebot auszeichnet. So sind bisher Daten aus über 80 Statistiken, über zum Teil lange Erhebungszeiträume, für die wissenschaftliche Forschung verfügbar. Das Datenangebot wird dabei ständig erweitert. Um für diesen großen Umfang an Daten eine zügige Bereitstellung für Nutzeranfragen oder die Erstellung ausführlicher Metadaten gewährleisten zu können, haben 13 statistische Landesämter die Zuständigkeit für einzelne Statistiken bzw. Statistikbereiche im FDZ übernommen.

Die Erarbeitung fachlicher Konzepte oder die Klärung rechtlicher Fragen erfolgt in enger Abstimmung mit dem FDZ des Statistischen Bundesamtes. Die gute Kooperation zwi-

18) Vgl. Hauser, R.; Wagner, G.; Zimmermann, F.: Memorandum: Erfolgsbedingungen empirischer Wirtschaftsforschung und empirisch gestützter wirtschafts- und sozialpolitischer Beratung – IZA Discussion Paper No. 14, Bonn, 1998, S. 13.

schen beiden FDZ zeigt sich aber auch bei der gemeinsamen Erstellung von Scientific-Use-Files oder CAMPUS-Files.

Mit den CAMPUS-Files haben die FDZ vollständig anonyme Datensätze für die statistische Methodenlehre in den Hochschulen entwickelt und damit Möglichkeiten für eine bessere Ausbildung der Studierenden geschaffen. Die Forderungen nach solchen Datensätzen wurden von der Wissenschaft bereits vor Einrichtung der KVI geäußert und fanden sich auch in dem vor 10 Jahren vorgelegten Gutachten wieder¹⁸⁾.

Verbesserung des Analysepotenzials durch Datenintegration

Einen innovativen Weg bei der Verbesserung der informationellen Infrastruktur verfolgen die FDZ seit einiger Zeit auch mit der Integration von Daten. Dabei werden Daten unterschiedlicher Statistiken miteinander kombiniert. Hierdurch besteht die Möglichkeit, für die Analy-

se vornehmlich ökonomischer Zusammenhänge Informationen aus verschiedenen Statistiken gemeinsam zu nutzen. Eine Datenintegration ist im Querschnitt, aber auch im Längsschnitt möglich. Durch die Zusammenführung von Einzeldaten verschiedener Erhebungswellen im Längsschnitt kann vor allem dem steigenden Bedarf nach Panel-Datensätzen nachgekommen werden.

Mit einer Integration von Datensätzen lässt sich der Analysegehalt erhobener Mikrodaten deutlich erhöhen. In der Folge können auch sehr komplexe Forschungsfragen beantwortet werden. Durch die Integration von Daten ist es aber auch möglich, zusätzliche statistische Erhebungen zu vermeiden und so nicht nur Kosten zu sparen, sondern vor allem die Auskunftgebenden vor zusätzlichen Belastungen zu bewahren. Die Integration von Daten wird daher nicht nur von der Wissenschaft gefordert, sondern sie wurde bei der Evaluation des FDZ der Statistischen Landesämter im Jahr 2006 von einer Gutachtergruppe auch ausdrücklich empfohlen¹⁹⁾. Mit dem 2005 neu gefassten § 13a Bundesstatistikgesetz hat der Gesetzgeber hierfür die rechtlichen Rahmenbedingungen geschaffen.

Blieb die Möglichkeit zur Integration von Daten in der Vergangenheit von der amtlichen Statistik weitestgehend ungenutzt, haben die FDZ der Statistischen Ämter des Bundes und der Länder diese Chance vor einiger Zeit ergriffen. Mit dem Projekt **Amtliche Firmendaten für Deutschland** (AFiD) des FDZ der statistischen Landesämter konnten Einzeldaten der Wirtschafts- und Umweltstatistiken im Längs- und Querschnitt über das Unternehmensregister integriert werden. Im Projekt **Kombinierte Firmendaten für Deutschland** (KombiFiD) werden erstmalig Datensätze erstellt, bei denen Erhebungen verschiedener Institutionen miteinander kombiniert werden. Neben Daten der amtlichen Statistik werden hier auch Informationen der Bundesagentur für Arbeit und der Deutschen Bundesbank verwendet. Insbesondere mit Blick auf die Erzeugung von Längsschnittdatensätzen ist an dieser Stelle auch das Projekt „Wirtschaftsstatistische Paneldaten und faktische Anonymisierung“ zu nennen.

Zukünftige Herausforderungen

Beide FDZ sind heute etablierte Schnittstellen zwischen Wissenschaft und amtlicher Statistik. Am 1. April 2012 wird das FDZ der statistischen Landesämter 10 Jahre

bestehen. Das FDZ des Statistischen Bundesamtes wird sein zehnjähriges Bestehen bereits im Herbst dieses Jahres feiern. Nach dem Rückblick auf die bisherige, sehr erfolgreiche Entwicklung der informationellen Infrastruktur erscheint an dieser Stelle ein Ausblick in die Zukunft angebracht. Als Dienstleistungseinrichtungen müssen sich die FDZ in hohem Maße an den Bedürfnissen ihrer Nutzer orientieren. Wie also könnte der Bedarf der Datennutzer in Zukunft aussehen und wo liegen zukünftige Herausforderungen für die FDZ der amtlichen Statistik?

Im Januar dieses Jahres fand in Wiesbaden die 5. Konferenz für Sozial- und Wirtschaftsdaten statt. Betrachtet man die hier in Vorträgen vorgestellten und in Foren diskutierten Themen, lassen sich ein paar Anhaltspunkte für die Beantwortung der aufgeworfenen Frage ableiten. Als eine Herausforderung darf sicherlich die zunehmende Nachfrage nach Geodaten bzw. die Georeferenzierung von Mikrodaten und die Tendenz zu kleinräumlichen Untersuchungen gesehen werden. Eine andere Herausforderung wird auch weiterhin die Verbesserung bestehender bzw. die Entwicklung neuer Zugangswege zu den Einzeldaten der amtlichen Statistik bleiben. Die Bedeutung der On-Site-Nutzung – also die Datenauswertung an einem Gastwissenschaftlerarbeitsplatz oder mittels kontrollierter Datenfernverarbeitung (KDFV) – ist mit einem Anteil von 75 % sehr hoch. Um insbesondere bei der KDFV die Bearbeitungszeiten so kurz wie möglich zu halten, wird derzeit in den FDZ der Statistischen Ämter des Bundes und der Länder an der Erstellung von syntaktisch und semantisch validen Strukturdatensätzen geforscht. Mit diesen verbesserten Strukturdatensätzen sollen die Wissenschaftler nicht nur testen können, ob Programme für die KDFV fehlerfrei ablaufen, sondern sie sollen vor allem auch Hinweise darauf erhalten, ob statistische Modelle zutreffen. Ein mögliches langfristiges Ergebnis dieser Forschung wäre die Erzeugung synthetischer Daten, die den Grundstein für ein „echtes“ Fernrechnen²⁰⁾ legen könnten.

Sei es die ständige Erweiterung des Datenangebotes um aktuelle Berichtsjahre, die Erstellung weiterer integrierter Datensätze oder die Bereitstellung von neuen Erhebungen, wie dem Zensus 2011, für den bereits jetzt bei der Wissenschaft erhebliche Nachfrage besteht, – der kontinuierliche Ausbau des Datenangebots wird bei den FDZ weiterhin eine entscheidende Rolle spielen.

Vor dem Hintergrund der hier genannten Punkte und mit Blick auf die immer komplexer werdenden Analysen durch die Wissenschaft wird die größte Herausforderung für die FDZ jedoch auch in Zukunft bei der Sicherstellung der statistischen Geheimhaltung liegen. Denn nur wenn diese garantiert ist, ist auch eine Weiterentwicklung der FDZ möglich.

19) Vgl.: Malchin, A.; Pohl, R.: Firmendaten der amtlichen Statistik – Datenzugang und neue Entwicklungen im Forschungsdatenzentrum, In: DIW Berlin (Hrsg.): Vierteljahreshefte zur Wirtschaftsforschung 76 (2007), S. 13 ff.

20) Unter „echtem“ Fernrechnen werden hier Möglichkeiten verstanden, bei denen die Wissenschaft vom eigenen Arbeitsplatz in der wissenschaftlichen Einrichtung aus über einen Server rund um die Uhr Analysen durchführen kann und binnen kurzer Zeit automatisch geprüfte und vollständig anonyme Ergebnisse erhält.

Fazit

Die FDZ sind heute eine unverzichtbare Schnittstelle zwischen Wissenschaft und öffentlichen Datenproduzenten. Diese Entwicklung war noch im Jahr 2001 kaum vorstellbar, denn es bestand ein deutliches Missverhältnis zwischen dem Bedarf der empirisch forschenden Wissenschaft nach gut dokumentierten Einzeldaten und entsprechenden Zugangsmöglichkeiten zu diesen Daten. Auf dieses Missverhältnis machte vor 10 Jahren ein Gutachten der KVI aufmerksam, das als ein Meilenstein auf dem Weg zur Einrichtung von FDZ bei den öffentlichen Datenproduzenten angesehen werden kann. Dieses

Gutachten hat auch den Weg für die FDZ der amtlichen Statistik bereitet.

Beide FDZ sind heute zweifellos ein enormer Erfolg. Sie müssen jedoch auch zukünftig den schwierigen Ausgleich zwischen den stetig wachsenden Anforderungen der Datennutzer und den bestehenden Geheimhaltungsbestimmungen leisten. Eine dauerhafte Finanzierung der FDZ ist noch immer nicht abschließend geklärt. Auch wenn heute, 10 Jahre nach Vorlage des Gutachtens, die FDZ der amtlichen Statistik von vielen als selbstverständlich angesehen werden, sind diese noch längst keine Selbstverständlichkeit.

Verzeichnis der Krankenhäuser, Vorsorge- und Rehabilitations- einrichtungen



STATISTIK HESSEN

Die vorliegende Publikation enthält Informationen zu allen Einrichtungen, die zur stationären Versorgung beigetragen haben (sofern ihre Zustimmung für eine Veröffentlichung vorlag); z. B. Anschriften (einschl. Telefon- und Faxnummern), Träger der Einrichtung, Fachabteilungen, jahresdurchschnittlich aufgestellte Betten. Der Nachweis erfolgt getrennt nach Krankenhäusern (gut 2200 Einrichtungen) und nach Vorsorge- oder Rehabilitationseinrichtungen (knapp 1400 Einrichtungen). Die Einrichtungen sind nach Ländern, Regierungsbezirken und Landkreisen sortiert.

Das Verzeichnis wird als ZIP-Datei über E-Mail versandt.

Aktuelle Angaben über die neuesten Ausgaben sowie Preise finden Sie im Internet oder nehmen Sie direkt Kontakt mit uns auf.

Zu beziehen direkt vom
Hessischen Statistischen Landesamt, 65175 Wiesbaden

Telefon: 0611 3802-951 · Fax: 0611 3802-992
E-Mail: vertrieb@statistik-hessen.de · Internet: www.statistik-hessen.de



W_158