

10 Jahre Forschungsdaten für die Wissenschaft - zum zehnjährigen Bestehen der Forschungsdatenzentren der amtlichen Statistik

Um der Wissenschaft einen komfortablen Zugang zu Einzeldaten der amtlichen Statistik zu ermöglichen, wurden vor über 10 Jahren Forschungsdatenzentren (FDZ) beim Statistischen Bundesamt und den statistischen Landesämtern eingerichtet. Aus diesem Anlass fand nun eine Konferenz statt, die sich sowohl rückblickend mit den FDZ beschäftigte, als auch einen Ausblick auf deren zukünftige Entwicklung zum Thema hatte. Der folgende Beitrag berichtet über diese Konferenz und zeichnet die Entwicklung beider FDZ der amtlichen Statistik in ihren Grundzügen nach.

Konferenz zum zehnjährigen Bestehen

Am 12. und 13. Juli feierten das Forschungsdatenzentrum des Statistischen Bundesamtes und das der statistischen Landesämter gemeinsam ihr zehnjähriges Bestehen. Zu diesem Anlass trafen sich Vertreter aus Wissenschaft und amtlicher Statistik in der hessischen Landesvertretung in Berlin. Die Veranstaltung bot beiden Seiten die Möglichkeit, sich miteinander über die Erfahrungen in den letzten 10 Jahren auszutauschen und Strategien für zukünftige Entwicklungen zu entwerfen. Unter den rund 170 Teilnehmern waren neben bekannten Wissenschaftlern aus der empirischen Sozial- und Wirtschaftsforschung auch viele Leiterinnen und Leiter der statistischen Ämter, so auch die Präsidentin des Hessischen Statistischen Landesamtes (HSL), Frau Dr. Christel Figgenger.

Begrüßt wurden die Teilnehmer durch den Präsidenten des Statistischen Bundesamtes, Herrn Roderich Egeler und die Präsidentin des Amtes für Statistik Berlin-Brandenburg, Frau Professor Ulrike Rockmann. Beide wiesen in ihren Ansprachen auf den schwierigen Weg hin, den die FDZ seit ihrer Gründung bestritten haben und weiterhin bestreiten. Sie würdigten dabei Personen, die sich insbesondere in der Aufbauphase der beiden FDZ um die informationelle Infrastruktur in Deutschland verdient gemacht haben. In diesem Zusammenhang wurde auch dem früheren Präsidenten des HSL, Herrn Eckart Hohmann, gedankt. Er hatte als langjähriger Vertreter der statistischen Landesämter im Rat für Sozial- und Wirtschaftsdaten die Einrichtung und Etablierung der FDZ entscheidend vorangetrieben und maßgeblich mitbeeinflusst.

Im Anschluss an die Begrüßung folgte eine Betrachtung der Entwicklung beider FDZ aus der Sicht der Wissenschaft. Hierbei wurde neben Lob auch Kritik geäußert. Während die Forschungsdatenzentren als solche sehr positiv bewertet wurden, kritisierten Wissenschaftler mehrfach die Kosten für den Datenzugang sowie die aus ihrer Sicht zu strikte Geheimhaltung in den statistischen Ämtern. Von Vertretern der amtlichen Statistik wurde hierzu angemerkt, dass die Wahrung des Datenschutzes keine Ermessenfrage ist, sondern auf Grundlage gesetzlicher Vorgaben geschieht. Der Schutz der Auskunftgebenden hat oberste Priorität. Nur die strikte Wahrung des Datenschutzes kann das Vertrauen der Auskunftgebenden in die amtliche Statistik sicherstellen und somit eine hohe Qualität der erhobenen Informationen garantieren. Dies kommt letztlich auch der Wissenschaft zugute. Bei der Diskussion um die Kosten zeigten die statistischen Ämter Verständnis für die Kritik. Hintergrund für die derzeitigen Kostensätze in den FDZ sind die Vorgaben von Bund und Ländern, wonach die FDZ sich aus den Einnahmen zur Hälfte selbst finanzieren sollen. Einigkeit bestand darin, dass es nicht zielführend sein kann, wenn aufgrund hoher Kosten Datennutzungen ausbleiben oder die Zahl der Datennutzer zurückgeht. Hohe Kosten, die sich als Hemmnis für den Datenzugang herausstellen, widersprechen der Idee, den Zugang der Wissenschaft zu amtlichen Mikrodaten zu erleichtern. Eine Idee, auf die sich die FDZ gründen.

Den Abschluss des ersten Veranstaltungstages bildete eine Poster-Präsentation, bei der Nachwuchswissenschaftler Arbeiten vorstellten, die auf der Grundlage amtlicher Mikrodaten durchgeführt wurden. Die einzelnen Beiträge wurden im Vorfeld von einer Kommission aus einer Vielzahl eingereicherter Bewerbungen ausgewählt. Unter den ausgewählten Projekten war auch die Arbeit von Lars Redert, der am FDZ-Standort Hessen mit Hilfe von Daten der hessischen Verkehrsunfallstatistik das Unfallgeschehen auf Autobahnen und Landstraßen untersuchte. Er prüfte dabei die Möglichkeiten zur Darstellung von Unfallschwerpunkten in webbasierten Kartendiensten mittels Georeferenzierung.

Während der erste Veranstaltungstag unter der Überschrift „Nutzen der Forschungsdatenzentren“ stand,



Lars Redert (HSL) im Gespräch mit Dr. Florian Köhler (Landesbetrieb für Statistik und Kommunikationstechnologie Niedersachsen). Foto: HSL

trug der zweite Tag den Titel „Forschungsdatenzentren in der Praxis“. Hier wurden Ergebnisse beispielhafter Datennutzungen aus verschiedenen Bereichen in 4 Teilveranstaltungen vorgestellt und diskutiert. Die ersten beiden Teilveranstaltungen umfassten Beiträge aus den Bereichen „Arbeitsmarkt“ und „Sozialstatistik“. Dabei wurden Untersuchungen zu Arbeitszeitmustern, zum Handlungsbedarf für eine soziale Absicherung Soloselbstständiger oder zu Einflussfaktoren bei der Inanspruchnahme von Elternzeit durch Väter präsentiert. Darüber hinaus wurden Hintergründe für die Abbrüche von Fremdpflegeverhältnissen bei Kindern und Jugendlichen, zur Lebenserwartung und Pflege im Alter sowie Qualitätsindikatoren der OECD auf der Grundlage der DRG-Statistik¹⁾ vorgestellt. Die dritte Teilveranstaltung fasste Forschungsergebnisse zusammen, die auf der Grundlage amtlicher Firmendaten entstanden sind. Hierzu zählten Analysen zur Effizienz von Krankenhäusern sowie die Betrachtung von Exporteinbrüchen oder das lebenslange Lernen in Unternehmen. In der

1) Die fallpauschalenbezogene Krankenhausstatistik wird verallgemeinert als DRG-Statistik bezeichnet. DRG steht dabei für Diagnosis Related Groups (dt: diagnosebezogene Fallgruppen). Die Statistik wird vom Statistischen Bundesamt durchgeführt.

vierten Teilveranstaltung stellten Wissenschaftler ihre Forschungsarbeiten zur Verbreitung von Tabakkonsum in Deutschland sowie zur permanenten Einkommenshypothese unter Liquiditätsbeschränkungen vor und gaben einen Überblick über Möglichkeiten zur räumlichen Auswertung amtlicher Statistikdaten. Nicht nur der zuletzt genannte Beitrag machte noch einmal deutlich, dass von vielen Forschungsprojekten auch die amtliche Statistik selbst profitieren kann.

Gründung und Aufgaben der Forschungsdatenzentren

Die Veranstaltung in Berlin war eine „nachträgliche Geburtstagsfeier“, denn das FDZ des Statistischen Bundesamtes wurde bereits am 1. September 2001 eingerichtet und das FDZ der statistischen Landesämter folgte am 1. April 2002. Die Einrichtung beider FDZ vor über 10 Jahren geschah mit dem Ziel, der wissenschaftlichen Forschung einen vereinfachten Zugang zu amtlichen Mikrodaten zu ermöglichen.

Mit derzeit 455 verschiedenen Statistiken sind die Statistischen Ämter des Bundes und der Länder mit Abstand die größten Datenproduzenten in Deutschland. Sei es die Zahl an Geborenen und Gestorbenen, die Zusam-

mensetzung der Steuereinnahmen von Bund, Ländern und Gemeinden, die Entwicklung am Arbeitsmarkt, die Veränderung ökonomischer Zusammenhänge oder die Häufigkeit von Hochschulabschlüssen – Daten der Statistikämter informieren über alle relevanten sozialen und wirtschaftlichen Zusammenhänge in der Bundesrepublik. Die von den statistischen Ämtern veröffentlichten Tabellen basieren in der Regel auf erhobenen Einzeldaten, und für diese Einzeldaten interessiert sich vor allem die Wissenschaft. Zum einen bedarf es für eine tiefgehende Beurteilung gesellschaftlicher und ökonomischer Zusammenhänge oftmals des Blickes „hinter“ einen zusammengefassten (Tabellen-)Wert. Dies gilt z. B. für die Beantwortung von Fragestellungen, die sich mit bestimmten Teilgruppen aus Wirtschaft und Gesellschaft beschäftigen²⁾. Zum anderen sind Einzeldaten eine Voraussetzung für die Anwendbarkeit vieler statistischer Methoden und Verfahren.

Auch wenn das Bundesstatistikgesetz mit dem sogenannten „Wissenschaftsprivileg“ den Zugang der Wissenschaft zu faktisch anonymen Mikrodaten für Forschungszwecke seit 1987 möglich macht, war vor der Einrichtung der FDZ der Datenzugang für die Wissenschaft sehr begrenzt³⁾. Von der amtlichen Statistik wurden für Forschungsprojekte zwar anonymisierte Einzeldaten zur Verfügung gestellt, allerdings war die Datenbereitstellung teilweise mit hohen Kosten verbunden und konzentrierte sich auf wenige Statistiken⁴⁾. Zumeist waren dies Statistiken, die im Zusammenhang mit haushalts- oder personenbezogenen Erhebungen standen. Für Einzeldaten aus den Wirtschaftsstatistiken, bei denen sich insbesondere bei tiefer räumlicher oder sachlicher Gliederung die faktische Anonymität nur schwer sicherstellen lässt, bot sich ein Datenzugang nur in geringem Umfang. In einigen wenigen Fällen hatten Wissenschaftler die Möglichkeit, auf Grundlage gesonderter Vereinbarungen mit einem statistischen Landesamt Einzeldaten innerhalb der geschützten Räume der amtlichen Statistik zu nutzen. Die Auswertung blieb da-

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

3) Vgl. Richter, Alexander, „10 Jahre Wege zu einer besseren informationellen Infrastruktur“, Staat und Wirtschaft in Hessen, 66. Jahrg., Heft 2/3 Februar/März 2011, S. 28 ff.

4) Ebd. S. 30.

5) 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.

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

7) Für detaillierte Informationen zum Gutachten siehe: Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (Hrsg.), „Wege zu einer besseren informationellen Infrastruktur“, Baden-Baden, 2001.

8) Vgl. Richter, Alexander u. Weil, Stefan, „Metadaten – Eine Grundlage für die Auswertung amtlicher Statistiken durch die Wissenschaft“, Statistische Monatshefte Rheinland-Pfalz, Heft 1/2005, S. 14 ff.

bei jedoch auf die Daten des jeweiligen Landesamtes beschränkt und war folglich nur für ein einzelnes Bundesland möglich⁵⁾.

Dies führte insbesondere bei den Wirtschaftswissenschaftlern dazu, dass die Forderungen nach Lösungen für die Beschränkungen immer lauter wurden. Viele Ökonomen sahen die internationale Wettbewerbsfähigkeit der empirischen Wirtschaftsforschung in Deutschland gefährdet. Im Jahr 1998 setzte sich daher u. a. der Wissenschaftsrat in einer Stellungnahme mit diesem Problem auseinander⁶⁾. Daraufhin wurde im Jahr 1999 vom Bundesministerium für Bildung und Forschung eine Kommission zur Verbesserung der informationellen Infrastruktur (KVI) eingesetzt. Sie führte eine umfassende Bestandsaufnahme der Situation durch und legte Vorschläge für eine Verbesserung vor. Das Ergebnis war ein im Jahr 2001 vorgelegtes Gutachten, das auf die Notwendigkeiten von erleichterten Zugangswegen zu Einzeldaten der amtlichen Statistik hinwies und die Einrichtung von Forschungsdatenzentren bei den amtlichen Datenproduzenten forderte⁷⁾.

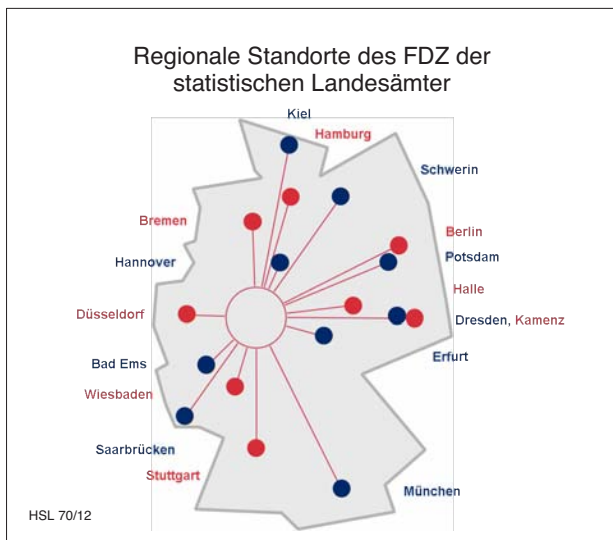
Die Tätigkeit der FDZ umfasst heute jedoch nicht allein die Aufbereitung und Bereitstellung von Mikrodaten für Forschungsprojekte. Um die Mikrodaten, die ausgewertet werden sollen, richtig interpretieren zu können, bedarf es Informationen, die eine korrekte Deutung von statistischen Daten ermöglichen, wie beispielsweise die Darstellung von Erhebungsmethoden oder Merkmalsdefinitionen⁸⁾. Eine Aufgabe der FDZ ist es daher ebenfalls, solche Zusatzinformationen – sogenannte Metadaten – für die wissenschaftlichen Nutzer zu erstellen und diesen zugänglich zu machen. Weitere Aufgaben sind die Wahrung der Datenschutzbestimmungen oder die Verbesserung des Analysepotenzials, beispielsweise durch Datenintegration.

Funktion der Forschungsdatenzentren

Das FDZ des Statistischen Bundesamtes und das FDZ der statistischen Landesämter sind heute etablierte Schnittstellen zwischen amtlicher Statistik und Wissenschaft.

Die gemeinsame Festveranstaltung machte noch einmal die von Beginn an bestehende, enge Zusammenarbeit





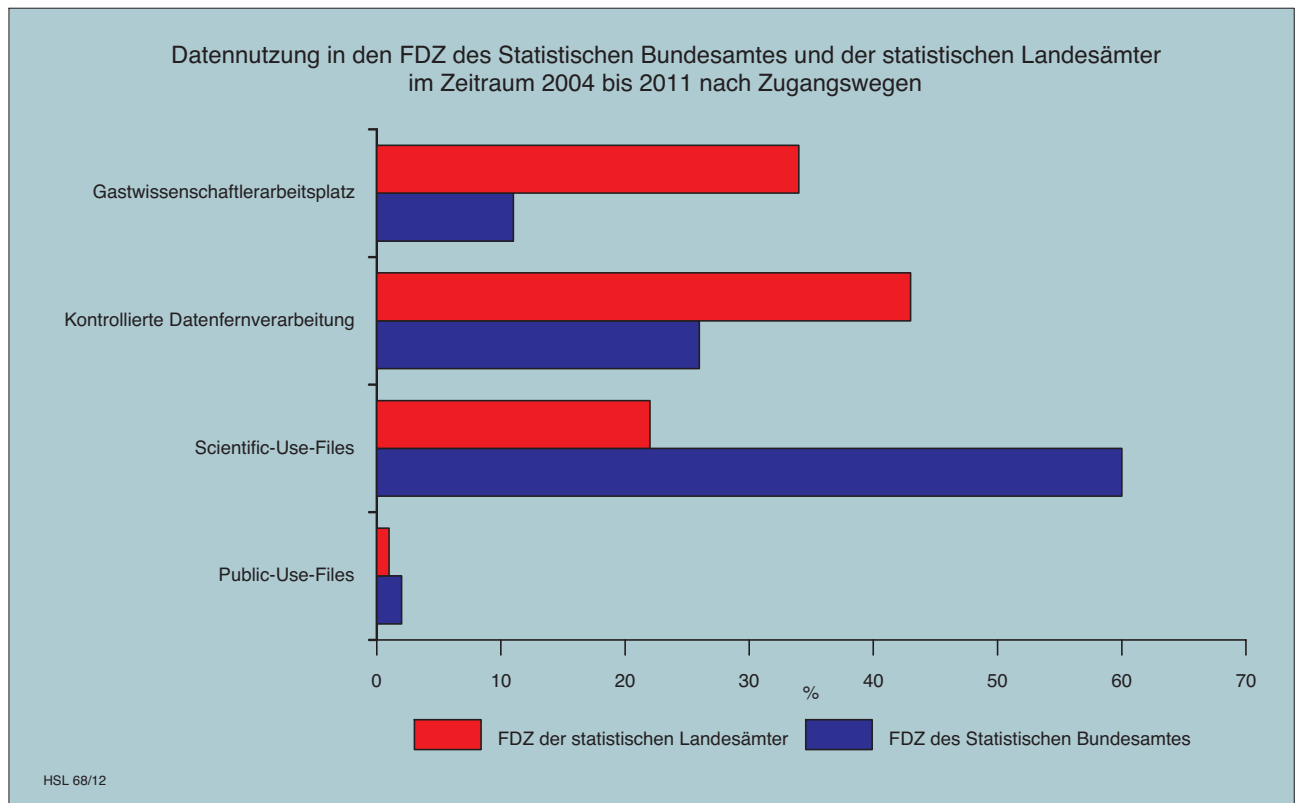
zwischen beiden FDZ deutlich. Den Zugang zu Mikrodaten der amtlichen Statistik ermöglichen beide. Es gelten dieselben rechtlichen und technischen Rahmenbedingungen. Trotzdem haben beide FDZ unterschiedliche Aufgaben- und Arbeitsschwerpunkte, die aufeinander abgestimmt sind. Während sich das FDZ des Statistischen Bundesamtes mit der Aufbereitung und Bereitstellung zentral erhobener Statistiken beschäftigt, ist das FDZ der statistischen Landesämter u. a. für dezentral durchgeführte Statistiken zuständig.

Das Forschungsdatenzentrum der statistischen Landesämter ist eine Arbeitsgemeinschaft der 14 statistischen Landesämter in Deutschland. Mit 18 Standorten bietet es eine bundesweite Präsenz. Insbesondere bei der Nut-

zung von Gastwissenschaftlerarbeitsplätzen ergibt sich aus dieser regionalen Struktur ein großer Vorteil für die Wissenschaft.

Die unterschiedliche regionale Präsenz der beiden FDZ führt zu einer weiteren Besonderheit. Während beim FDZ des Statistischen Bundesamtes mit rund 60 % der Datennutzungen ein Schwerpunkt im Vertrieb von Scientific-Use-Files für die Off-Site-Nutzung liegt, dominiert dagegen beim FDZ der statistischen Landesämter die sogenannte On-Site-Nutzung. Hierbei werden die Daten in einem statistischen Amt ausgewertet. Beim FDZ der statistischen Landesämter finden 77 % der Datennutzungen entweder an einem Gastwissenschaftlerarbeitsplatz in einem der 18 Standorte oder per kontrollierter Datenfernverarbeitung statt.

Die beiden FDZ haben ihre Arbeitsschwerpunkte aufeinander abgestimmt, um dem enormen Datenbedarf der empirisch arbeitenden Wissenschaft gerecht zu werden. Damit aber auch der große Umfang an Statistiken im Zuständigkeitsbereich der statistischen Landesämter adäquat bearbeitet werden kann, wurden darüber hinaus beim FDZ der statistischen Landesämter die fachlichen Zuständigkeiten zwischen den regionalen Standorten aufgeteilt. Im Ergebnis ist ein regionaler Standort, stellvertretend für alle 16 Bundesländer, für einen oder mehrere Statistikbereiche fachlich zuständig. Das heißt, dass ein Standort die Mikrodaten bestimmter Statistiken, die für Forschungszwecke genutzt werden können, für alle 16 Bundesländer zentral vorhält, diese für Datennutzungen im gesamten Bundesgebiet aufbereitet und die Metadaten erstellt.



FDZ-Standort Hessen

10 Jahre Forschungsdaten für die Wissenschaft bedeuten auch 10 Jahre FDZ-Standort Hessen. Der Standort im HSL ist seit seinem Bestehen für die Bereiche „Löhne und Gehälter“, „Erwerbstätigkeit“, „Betriebliche Weiterbildung“ sowie „Preise“ zuständig. Im Sommer 2007 wurde zusätzlich die fachliche Zuständigkeit für den komplexen Bereich der Unternehmensteuerstatistiken übernommen. Seitdem werden in Hessen auch die Statistik der Umsatzsteuervoranmeldungen und -veranlagungen, die Gewerbesteuerstatistik sowie die Körperschaftsteuerstatistik für wissenschaftliche Anfragen betreut.

Seit 2011 wird im Rahmen eines Pilotprojektes mit dem Hessischen Kultusministerium für eine derzeit begrenzte Anzahl an Forschungsprojekten auch der Zugang zu Daten der hessischen Schulstatistik⁹⁾ am Gastwissenschaftlerarbeitsplatz in Wiesbaden ermöglicht. Schuldaten sind „Ländersache“. Die Datenlage ist hier, anders als bei anderen amtlichen Statistiken, bundesweit uneinheitlich. So gibt es in vielen Bundesländern keine Einzeldaten in diesem Bereich. Mit dem Pilotprojekt versuchen das Hessische Kultusministerium und das HSL dem Bedarf der Wissenschaft nach schulstatistischen Einzeldaten nachzukommen.

Insbesondere die fachliche Zuständigkeit für die Unternehmensteuerstatistiken fügt sich gut in ein weiteres Projekt ein, mit dem der Standort Hessen zusammen mit dem FDZ des Statistischen Bundesamtes die bereits gute informationelle Infrastruktur im Rhein-Main-Gebiet ausbauen wird. Im Herbst sollen im Rahmen eines Gemeinschaftsprojektes des HSL und des Statistischen Bundesamtes Gastwissenschaftlerarbeitsplätze im House of Finance (HoF) an der Goethe-Universität Frankfurt eingerichtet werden. Das HoF ist eine einzigartige Forschungs- und Lehreinrichtung für Rechts- und Wirtschaftswissenschaften. Zum Kuratorium des HoF gehören neben dem Bundesfinanzminister und dem Bundesbankpräsidenten auch der hessische Ministerpräsident sowie Vertreter internationaler Finanzinstitute.

Fazit

Beide Forschungsdatenzentren blicken auf 10 Jahre erfolgreiche Arbeit zurück. Sie müssen jedoch zukünftig den immer schwieriger werdenden Ausgleich zwischen

den stetig wachsenden Anforderungen der Datennutzer auf der einen Seite und den bestehenden Geheimhaltungsbestimmungen sowie begrenzten Ressourcen auf der anderen Seite meistern.

Das HSL gestaltete die Entwicklung der informationellen Infrastruktur in Deutschland von Anfang an aktiv mit. So vertrat der frühere Präsident des HSL die statistischen Landesämter seit Gründung im Rat für Sozial- und Wirtschaftsdaten. Das HSL ist seit Bestehen der FDZ an einer Vielzahl von Arbeitsgruppen und Methodenprojekten, die mit der Forschungsdateninfrastruktur im Zusammenhang stehen, beteiligt. Insofern war es naheliegend, dass der Festakt zum zehnjährigen Bestehen der Forschungsdatenzentren in Berlin in der hessischen Landesvertretung stattfand.

Was ist ...

... der Rat für Sozial- und Wirtschaftsdaten?

Der Rat für Sozial- und Wirtschaftsdaten (RatSWD) hat die Aufgabe, die Dateninfrastruktur in Deutschland zu bewerten, weiterzuentwickeln sowie für Bund und Länder Empfehlungen zum Programm der wissenschaftsgetragenen Statistik und dessen Finanzierung zu erarbeiten. Dies geschieht mit dem Ziel, die Situation der empirisch arbeitenden Sozial- und Wirtschaftswissenschaften nachhaltig zu verbessern und zu ihrer internationalen Wettbewerbsfähigkeit beizutragen. 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. Der RatSWD setzt sich sowohl aus Vertretern der Wissenschaft als auch der Datenproduzenten zusammen. (<http://www.ratswd.de>)

... der Wissenschaftsrat?

Der Wissenschaftsrat (WR) wurde 1957 von Bund und Ländern eingerichtet, um einen Überblick über die wissenschaftliche Arbeit in Deutschland zu geben und die Bundesregierung bzw. die 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 Bundesregierung und der Landesregierungen zusammen (<http://www.wissenschaftsrat.de>).

⁹⁾ Die hessische Schulstatistik umfasst neben Informationen zu Schulen auch Informationen zu Schülern, Lehrern und Unterricht in Hessen.