



Datenverknüpfung für Statistik und Forschung – ein Werkstattgespräch

Am 15. Februar hat das Bundesministerium für Wirtschaft und Klimaschutz (BMWK) gemeinsam mit dem Bundesministerium für Bildung und Forschung (BMBF) ein Werkstattgespräch zum Thema Verknüpfung von (Personen-)Daten für Statistik und Forschung durchgeführt. Ziel war es, im Kontext der Erarbeitung des Forschungsdatengesetzes, für das Thema fehlende Möglichkeiten der Datenverknüpfung zu sensibilisieren. Die Verknüpfung von Daten ist einerseits häufig Voraussetzung dafür, gute Politikberatung und Forschungsergebnisse zu erzielen. Andererseits lassen sich Statistiken durch Datenverknüpfungen belastungsärmer erheben oder werden hierdurch überhaupt erst ermöglicht.

Das Gespräch fand im Kontext der Arbeiten der Interministeriellen Arbeitsgruppe „Bessere Wirtschaftsstatistik und erleichterter Datenzugang für die Wissenschaft“ statt. Die Arbeitsgruppe aus BMBF, BMF, BMI, StBA und BMWK hatte sich vor rund zwei Jahren auf ein Eckpunktepapier verständigt, welches zahlreiche Vorschläge zur Verbesserung der Wirtschaftsstatistik und des Datenzugangs für die Forschung enthielt. Auch das Thema Verknüpfung von (Personen-)Daten für Statistik und Forschung wurde darin aufgegriffen. Mit den Vereinbarungen des Koalitionsvertrags

zur 20. Legislaturperiode, insbesondere zum Forschungsdatengesetz (FDG), sollen auch die Forderungen des Eckpunktepapiers umgesetzt werden. Darauf hatte sich Bundesminister Habeck (BMWK) mit den Bundesministerinnen Faeser (BMI) und Stark-Watzinger (BMBF) verständigt.

Am Werkstattgespräch nahmen zahlreiche Sachverständige aus Wissenschaft und einschlägigen Institutionen und Gremien mit aktiven Beiträgen teil (siehe weiter unten). Zusätzlich waren der Vorsitzende der Monopolkommission Prof. Jürgen Kühling und Prof. Simon Jäger vom MIT als Sachverständige eingeladen. Außerdem waren Vertreterinnen und Vertreter zahlreicher Bundesministerien (darunter BKAm, BMF, BMI, BMJ, BMAS, BMG und andere), des Bundesbeauftragten für den Datenschutz und die Informationsfreiheit (BfDI) und des Statistischen Bundesamts (StBA) anwesend.

Die Teilnehmenden diskutierten Best Practices sowie Herausforderungen und Verbesserungsvorschläge bei der Datenverknüpfung. Alle Vortragenden hoben hervor, dass fehlende Verknüpfungsmöglichkeiten bei (Personen-)Daten erfolgversprechende Forschungsarbeiten behinderten. Wichtige wirtschaftspolitische Fragen in Deutschland könnten ohne Datenverknüpfung



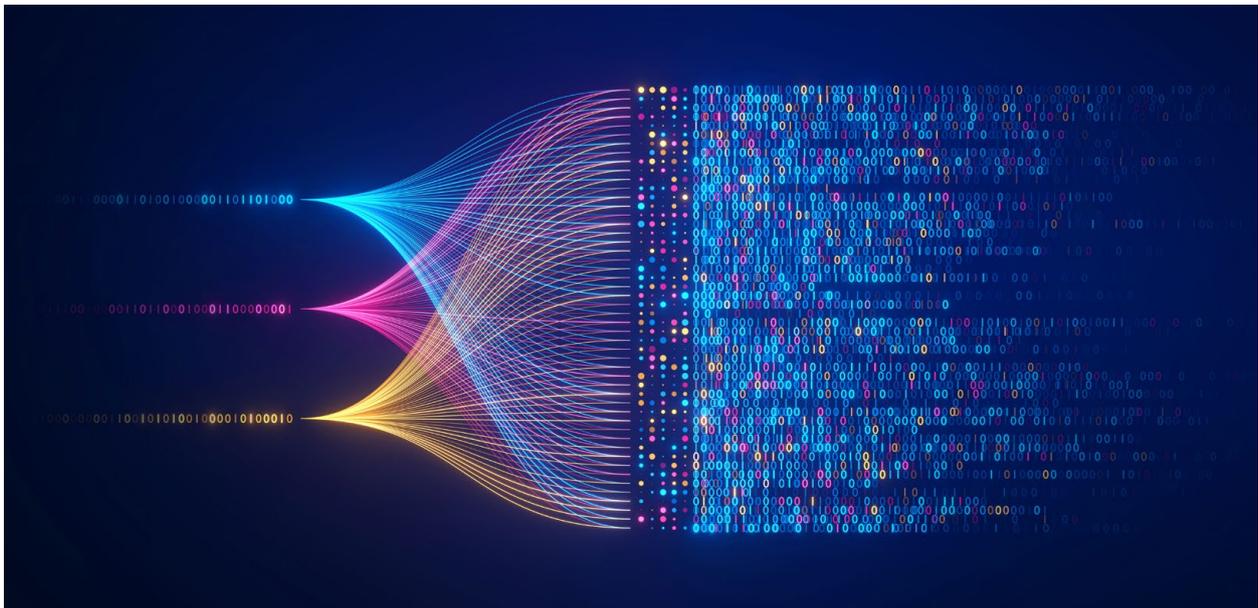
oft nicht evidenzbasiert beantwortet werden. Außerdem würden die Weiterentwicklung und Durchführung von Statistiken erschwert oder unmöglich gemacht. Deutschland habe teilweise ideale (Verwaltungs-)Daten, die aber aufgrund von Restriktionen häufig nicht genutzt oder mit anderen Daten sinnvoll verknüpft werden könnten. In der Folge würden manche Projekte gar nicht, nicht vollständig oder mit ausländischen Daten, zum Beispiel aus skandinavischen Ländern, Österreich, den Niederlanden oder Portugal durchgeführt. Insoweit existierten für die Forschenden in begrenztem Umfang Ausweichmöglichkeiten. Allerdings sei die Übertragbarkeit der Forschungsergebnisse auf Deutschland häufig nicht gegeben, so dass die benötigte wirtschaftspolitische Evidenz für Deutschland dann nicht vorhanden sei.

DIE POSITIONEN WICHTIGER STAKEHOLDER

Dr. Ruth Brand, Präsidentin des Statistischen Bundesamtes (StBA), zeigte sich überzeugt, dass eine Erweiterung der Verknüpfungsmöglichkeiten zu besseren Ergebnissen sowie einer belastungsärmeren und effizienteren Statistikproduktion führen würde und dass damit neue Daten für die Forschung gewonnen und die amtliche Statistik weiterentwickelt werden könnten. Zudem sei ein gesetzlich festgeschriebener Forschungsauftrag für das Statistische Bundesamt (StBA) ein wichtiger Schritt, um das StBA stärker für die Wissenschaft zu öffnen. Der Datenzugang für die Forschung solle durch ein beim Statistischen Bundesamt angesiedeltes German Micro Data Center verbessert werden. Dort könnten Daten der amtlichen Statistik unter sicheren Bedingungen mit Daten anderer Anbieter verknüpft werden – mit dem Ziel, einen vereinfachten Datenzugang für die Wissenschaft und verbesserte datenbasierte Forschung zu ermöglichen.

Frau Prof. Riphahn, Vorsitzende des Vereins für Socialpolitik (VfS), berichtete über die Ergebnisse einer Mitgliederbefragung des VfS, die die Unzufriedenheit der Befragten mit der Datensituation in Deutschland verdeutlicht. Nach Ansicht der Befragten leide unter der unzureichenden Datensituation auch die Politikberatung. Frau Riphahn berichtete von politikrelevanten Forschungsprojekten in den Niederlanden, Schweden, Frankreich und Dänemark, die in Deutschland aufgrund fehlender Verknüpfungsmöglichkeiten nicht umsetzbar seien. In Österreich, den Niederlanden, Frankreich, Luxemburg, der Schweiz und Dänemark seien die vorhandenen Möglichkeiten zur Datenverknüpfung hingegen vorbildlich, obwohl fast alle der genannten Länder auch der Datenschutzgrundverordnung (DSGVO) und somit demselben europäischen Datenschutzrahmen wie Deutschland unterlägen. Laut Prof. Riphahn würde die DSGVO in Deutschland häufig zu restriktiv ausgelegt. Sie argumentierte, dass Gesetzesänderungen benötigt würden, die Verknüpfungen beispielsweise in einem German Micro Data Center (GMDC) als Datentreuhänder ermöglichen. Zudem bedürfe es eindeutiger Identifikatoren, der Datenschutz des Bundes und der Bundesländer solle harmonisiert werden und es solle eine Datenzulieferungspflicht öffentlicher Stellen an das GMDC geben.

Frau Prof. Schneider argumentierte als stellvertretende Vorsitzende des Rats für Sozial- und Wirtschaftsdaten (RatSWD), der insgesamt 42 akkreditierte Forschungsdatenzentren (FDZ) in Deutschland betreut, dass Deutschland über ein Netz hervorragender FDZ verfüge, diese aber aufgrund rechtlicher Beschränkungen häufig nicht miteinander verbunden werden könnten. Die Kernforderung des RatSWD zum Forschungsdatengesetz laute daher, diese „Datensilos“ über zusätzliche Verknüpfungsmöglichkeiten zu überbrücken. 





fungsmöglichkeiten zu einer Datenlandschaft zu verbinden. Frau Schneider bemängelte außerdem fehlende Remote-Access-Möglichkeiten (Datenfernverarbeitung über sichere Verbindungen), Datenlücken, etwa in den Bereichen Bildung und Gesundheit, fehlende Identifikatoren und zu restriktive Löschpflichten.

Herr Dr. Garnadt, Generalsekretär des Sachverständigenrats Wirtschaft (SVR), erläuterte die Sicht des SVR, der dieses Jahr dem Thema Forschungsdateninfrastruktur ein eigenes Kapitel im Jahresgutachten gewidmet hat. Als Gründe für fehlende Verknüpfungsmöglichkeiten nannte Herr Garnadt unter anderem das Bundesstatistikgesetz, welches zwar die Verknüpfung von Unternehmensdaten zwischen dem Statistischen Verbund und der Deutschen Bundesbank in Einzelfällen ermögliche, nicht aber die Verknüpfung mit den Betriebs- oder Beschäftigtendaten der Bundesagentur für Arbeit. Dabei könnten Erhebungsaufwände auf Seiten der Statistik und der Befragten reduziert werden, wenn Verknüpfungen ermöglicht würden. Zugleich würden Analyse-möglichkeiten hierdurch erheblich erweitert. Als konkrete Beispiele führte Garnadt die Mindestlohnevaluation und die Messung der Auswirkungen der aktuellen geopolitischen Spannungen auf Unternehmen und Beschäftigte an.

FALLBEISPIELE AUS DER FORSCHUNG

Herr Prof. Peichl, Leiter des Forschungszentrums Makroökonomik und Befragungen des ifo Instituts, berichtete über das 2023 gegründete Netzwerk für empirische Steuerforschung (NeSt), das seine Aufgaben nur mit zusätzlichen Verknüpfungsmöglichkeiten überzeugend erfüllen könne.

Herr Peichl nannte Verknüpfungen, die für Forschungszwecke wichtig seien, darunter beispielsweise die vertikale Verknüpfung von Unternehmensstatistiken (zum Beispiel Umsatzsteuerstatistik und Statistik über die Personengesellschaften) oder die Verknüpfung von Arbeitsmarkt- und Sozialversicherungsdaten mit Daten aus der Lohn- und Einkommensteuerstatistik.

Frau Prof. Dwenger von der Universität Hohenheim berichtete von ihrem Forschungsprojekt, bei dem Wirkungen von Reformen beim Rentenbezug für Frauen auf deren Arbeitsangebot im Erwerbsleben untersucht werden sollten. Hierzu sei das Heiratsdatum eine entscheidende Information. Amtliche Daten aus Deutschland dürften nach geltender Rechtslage jedoch nicht genutzt werden, sodass auf schwedische Daten zurückgegriffen werden müsse. Aufgrund abweichender sozialer Normen in Schweden würde dies erwartungsgemäß jedoch zu Ergebnissen führen, die nicht ohne weiteres auf die deutsche Gesellschaft übertragen werden könnten.

Mit einer ähnlichen Problematik sieht sich Frau Prof. Grumpert von der Universität Tübingen bei ihrem durch den European Research Council geförderten Forschungsprojekt zu Verlagerungen der Produktion von Gütern und Dienstleistungen ins Ausland konfrontiert. So genanntes Offshoring wirke sich auch auf den heimischen Arbeitsmarkt aus. Befürchtungen, dass dies zu einem Rückgang der Beschäftigung in Deutschland führe, hätten sich in der Vergangenheit allerdings nicht immer als berechtigt erwiesen. Neue Daten, die im Auftrag des BMWK beim StBA erstellt wurden, bestätigten dies. Allerdings



verschöben sich mit einer Verlagerung von Arbeitsplätzen ins Ausland Wertschöpfungsstrukturen im Inland und damit wahrscheinlich auch Qualifikationsstrukturen der Arbeitsnachfrage. Um die Auswirkungen von Offshoring auf die Beschäftigungsstrukturen zu untersuchen, würden Informationen über Beschäftigte benötigt. Diese liegen beim IAB vor und könnten technisch mit den Außenhandelsdaten beim StBA verknüpft werden. Allerdings ist die benötigte Verknüpfung nach geltender Rechtslage derzeit nicht möglich. Daher müsse die Frage mit norwegischen Daten untersucht werden, wo aber aufgrund der abweichenden Außenhandelsstruktur andere Ergebnisse erwartet würden, die ebenfalls nicht ohne weiteres auf die deutsche Wirtschaft übertragbar seien.

Zudem präsentierte Frau Gumpert eine Liste von fünf weiteren Forschungsprojekten, die allesamt ebenfalls an der Verknüpfung von Daten scheiterten. Die Projekte würden daher teilweise mit portugiesischen Daten sowie mit Daten aus der Schweiz durchgeführt.

WIE KANN FORSCHUNGS- UND STATISTIKFREUNDLICHER DATENZUGANG GEREGLT WERDEN?

Herr Prof. Fitzenberger, Direktor des Instituts für Arbeitsmarkt- und Berufsforschung (IAB), erläuterte, dass das IAB seit langem aufgrund der Regelungen der Sozialgesetzbücher (insbesondere §75, SGB X) Daten erfolgreich verknüpfe und daraus vielfältige Datenprodukte generiere, die der Forschung angeboten würden. Auf der Grundlage dieser Daten würde einerseits viel erfolgreiche Forschung durchgeführt und andererseits erfolgreich evidenzbasierte Politikberatung betrieben. Daher seien die Regelungen des SGB X teilweise als vorbildhaft anzusehen. Zusätzlich bestehe die Verknüpfungsmöglichkeit von Verwaltungsdaten mit Umfragedaten, wenn eine Zustimmung der Befragten vorliege. Allerdings seien Analysen auf die Arbeitsmarkt-, Berufs- und Sozialpolitikforschung eingengt und es fehlten Verknüpfungsmöglichkeiten mit anderen Rechtskreisen. Dies führe dazu, dass Lebensverlaufsdaten, die Bildung, Ausbildung und Arbeitsmarktergebnisse verknüpfen, in Deutschland rar seien. Zudem seien Verknüpfungen mit Daten anderer Datenproduzenten (z. B. Statistisches Bundesamt, Deutsche Bundesbank, Deutsche Rentenversicherung) oft rechtlich nicht durchführbar. Weiter gebe es nur wenige Verwaltungsdaten, die den Haushaltszusammenhang erfassen, und fast keine Verwaltungsdaten, die verschiedene Generationen verknüpfen. Aufgrund fehlender Verknüpfungsmöglichkeiten entstünden beispielsweise massive Defizite in der Mindestlohnforschung, beim Übergang von der Schule zum Beruf und in Bezug auf die Kindergrundsicherung. Schließlich stelle das Erfordernis der Einwilligung der Befragten zu der Verknüpfung von Um-

fragedaten und Verwaltungsdaten für bestimmte wichtige Forschungsfragen ein großes Hindernis dar, da Nicht Einwilligung zu Verzerrung in Forschungsergebnissen führen könne.

Frau Prof. Specht-Riemenschneider, Vorsitzende der Forschungsstelle Datenrecht, argumentierte, dass nicht das Datenschutzrecht die Datenverarbeitung und damit auch die Datenverknüpfung zu Forschungszwecken behindere, sondern das vergleichsweise restriktiv gestaltete Bundesstatistikgesetz (BStatG). Das BStatG könne im Einklang mit Art. 2 I, 1 I Grundgesetz und Art. 7 und 8 Grundrechtecharta der Europäischen Union forschungsfreundlicher ausgestaltet werden, wenn zugleich technische Sicherungsmechanismen die Verhältnismäßigkeit der Normen sicherstellten. Insbesondere könnten nach ihrer Auffassung die Paragraphen 5 Anordnung von Bundesstatistiken, 5a Nutzung von Verwaltungsdaten, 7 Erhebung für besondere Zwecke und 13a Zusammenführung von Daten des BStatG forschungsfreundlicher ausgestaltet werden. Die Zusammenführung von Daten zu statistischen Zwecken, aber auch von nicht zu statistischen Zwecken erhobenen Daten ließe sich auf Grundlage von Art. 6 Abs. 4 DSGVO in Verbindung mit einer entsprechenden nationalen Regelung realisieren. Eine technische Sicherung sollte über Datentreuhandmodelle und über das sog. „5 safes framework“ (safe projects, safe people, safe data, safe settings, safe outputs) gewährleistet werden.

ABSCHLUSSDISKUSSION

In der abschließenden Diskussion wurde über notwendige und geplante Änderungen am Bundesstatistikgesetz im Zuge des Forschungsdatengesetzes gesprochen. Es wurde betont, dass außer dem BStatG weitere Gesetze in den Blick genommen werden müssten. Darüber hinaus wurde auch aus dem Kreis der anwesenden Ressorts auf die Notwendigkeit bereichsübergreifender Verknüpfungen von Daten wie beispielsweise von Gesundheitsdaten und Sozialdaten für Forschungszwecke hingewiesen. Kritische Argumente zu einer besseren Verknüpfbarkeit von Daten etwa im Rahmen des anstehenden Forschungsdatengesetzes wurden nicht vorgebracht. Weitere Themen waren die Zersplitterung der Datenschutzaufsicht durch Bundeslandzuständigkeiten und die Frage, ob es bei der Verknüpfung von Daten eines eindeutigen Identifikators bedürfe. —

KONTAKT & MEHR ZUM THEMA

Referat: IC3 – Wachstum, Demografie, Statistik
schlaglichter@bmwk.bund.de