

Hertweck, Friederike; Isphording, Ingo Eduard; Matthewes, Sönke Hendrik;
Schneider, Kerstin; Spieß, C. Katharina

Working Paper

Für einen besseren Datenzugang im Bereich Bildungsdaten: Stellungnahme aus dem Verein für Socialpolitik

RatSWD Working Paper, No. 282

Provided in Cooperation with:

German Data Forum (RatSWD)

Suggested Citation: Hertweck, Friederike; Isphording, Ingo Eduard; Matthewes, Sönke Hendrik;
Schneider, Kerstin; Spieß, C. Katharina (2023) : Für einen besseren Datenzugang im Bereich
Bildungsdaten: Stellungnahme aus dem Verein für Socialpolitik, RatSWD Working Paper, No. 282,
Rat für Sozial- und Wirtschaftsdaten (RatSWD), Berlin,
<https://doi.org/10.17620/02671.87>

This Version is available at:

<https://hdl.handle.net/10419/283928>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen
Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle
Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich
machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen
(insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten,
gelten abweichend von diesen Nutzungsbedingungen die in der dort
genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

*Documents in EconStor may be saved and copied for your personal
and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to
exhibit the documents publicly, to make them publicly available on the
internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content
Licence (especially Creative Commons Licences), you may exercise
further usage rights as specified in the indicated licence.*



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

RatSWD Working Paper Series

RatSWD.
Rat für Sozial- und
WirtschaftsDaten

282

Für einen besseren Datenzugang im Bereich Bildungsdaten

Stellungnahme aus dem
Verein für Socialpolitik

Friederike Hertweck, Ingo E. Isphording,
Sönke H. Mathewes, Kerstin Schneider,
C. Katharina Spieß

Dezember 2023

www.ratswd.de

RatSWD Working Papers

des Rates für Sozial- und Wirtschaftsdaten (RatSWD)

Die *RatSWD Working Papers*-Reihe startete Ende 2007. In dieser Online-Publikationsreihe werden konzeptionelle und historische Arbeiten, die sich mit der Gestaltung der statistischen Infrastruktur und der Forschungsinfrastruktur in den Sozial-, Verhaltens- und Wirtschaftswissenschaften beschäftigen, publiziert. Dies sind insbesondere Papiere zur Gestaltung der Amtlichen Statistik, der Ressortforschung und der akademisch getragenen Forschungsinfrastruktur sowie Beiträge, die Arbeit des RatSWD selbst betreffend. Auch Papiere, die sich auf die oben genannten Bereiche außerhalb Deutschlands und auf supranationale Aspekte beziehen, sind besonders willkommen.

RatSWD Working Papers sind nicht-exklusiv, d. h. einer Veröffentlichung an anderen Orten steht nichts im Wege. Alle Arbeiten können und sollen auch in fachlich, institutionell und örtlich spezialisierten Reihen erscheinen.

Die Inhalte der *RatSWD Working Papers* stellen ausdrücklich die Meinung der jeweiligen Autorinnen bzw. Autoren dar und nicht die des RatSWD. Die Zuwendungsgeber des RatSWD haben die Publikationen nicht beeinflusst.

Herausgeberin oder Herausgeber der RatSWD Working Papers-Reihe ist die/der Vorsitzende des RatSWD:

seit 2020 Monika Jungbauer-Gans

2014–2020 Regina T. Riphahn

2009–2014 Gert G. Wagner

2007–2008 Heike Solga

Für einen besseren Datenzugang im Bereich Bildung

Stellungnahme aus dem Verein für Socialpolitik

Friederike Hertweck (RWI), Ingo E. Isphording (IZA),
Sönke H. Mathewes (Universität Potsdam),
Kerstin Schneider (Bergische Universität Wuppertal),
C. Katharina Spieß (BiB und Universität Mainz)

Dezember 2023

<https://doi.org/10.17620/02671.87>

Dieses Paper ist außerdem beim Verein für Socialpolitik erschienen.:
https://www.socialpolitik.de/de/vfs_stellungnahmen_daten_bildung

Inhalt

1. Einleitung.....	4
2. Fehlende Daten und Politikevaluation.....	6
3. Datenbasis für ein Bildungsverlaufsregister.....	9
4. Keine Utopie, sondern europäische Realität.....	14
5. Konkrete Schritte und Umsetzungsempfehlungen.....	17

1. Einleitung

Diese Stellungnahme fordert die Umsetzung eines **Bildungsverlaufsregisters über alle Stufen formaler Bildung**, um eine gravierende Datenlücke in Deutschland zu schließen, und sie fordert den **regelmäßigen Zugang der Forschung** zu diesen Daten. Dieses Bildungsverlaufsregister soll perspektivisch mit weiteren Datenquellen beispielsweise aus den Bereichen Arbeitsmarkt, Steuern und Gesundheit verknüpfbar sein, um auch **langfristige Effekte** von Bildung analysieren zu können. International würde Deutschland mit einem derartig verknüpfbaren Bildungsverlaufsregister an die Standards anderer europäischer Staaten anschließen.

Hintergrund: Bessere Daten, bessere Bildung

Bildung ist von zentraler Bedeutung für die wirtschaftliche und gesellschaftliche Entwicklung und damit den Wohlstand Deutschlands. Bildung ist auch Voraussetzung für die Teilhabe und Entwicklung jedes einzelnen Menschen. Es ist die Aufgabe der Politik, das Bildungssystem so zu gestalten, dass es diesen Anforderungen gerecht wird – gerade in einer sich schnell verändernden Gesellschaft. Insbesondere die demografische Entwicklung, die ökologische Transformation, Migrationsströme, der technologische Wandel sowie der schwindende gesellschaftliche Zusammenhalt fordern das Bildungssystem und die Bildungspolitik besonders heraus. Um das Bildungssystem angesichts dieser Herausforderungen weiterzuentwickeln, ist empirische Evidenz über das Funktionieren des Bildungssystems und die Wirkung von bildungspolitischen Maßnahmen von zentraler Bedeutung. Der Auftrag zum Bildungsmonitoring hat mit Artikel 91b GG Satz 2¹ Eingang in das Grundgesetz gefunden.

Doch trotz der hohen und unstrittigen Bedeutung von empirischer Evidenz für gute Bildungspolitik ist die Datenbasis für die Bildungsforschung in Deutschland unzureichend. In einer Online-Befragung der VfS-Mitglieder Anfang 2023 stimmten 76 % der Teilnehmenden mit dem Arbeitsgebiet Bildungsökonomik der Aussage zu, dass es der mangelnde Datenzugang erschwert, relevante und aktuelle Politikberatung zu fundieren.²

¹ Artikel 91b GG Satz 2: „(2) Bund und Länder können auf Grund von Vereinbarungen zur Feststellung der Leistungsfähigkeit des Bildungswesens im internationalen Vergleich und bei diesbezüglichen Berichten und Empfehlungen zusammenwirken.“

² Der Verein für Socialpolitik (VfS) hat zu Beginn des Jahres 2023 eine Online-Befragung seiner Mitglieder zur Zufriedenheit mit dem Zugang zu Forschungsdaten durchgeführt. Von den über 500 Teilnehmenden gaben 63 das Arbeitsgebiet Bildungsökonomik an. Ein zentrales Ergebnis der Befragung ist eine weitverbreitete Unzufriedenheit mit dem Datenzugang in Deutschland unter Forschenden in diesem Arbeitsgebiet. 59 % der Befragten sind eher unzufrieden oder sehr unzufrieden mit dem Datenzugang. 71 % geben an, dass Projekte wegen mangelnder Daten nicht umgesetzt werden konnten. 49 % der Forschenden arbeiten mit nicht-deutschen Daten, weil geeignete deutsche Daten nicht verfügbar sind. 82 % der Befragten stimmen der Aussage zu, dass der Datenzugang in Deutschland schlechter ist als in vergleichbaren Ländern und 67 % glauben nicht, dass es die Datenlage erlaubt, im internationalen Forschungswettbewerb mithalten zu können. Der Aussage, dass es der Zugang zu Daten erschwert, relevante und aktuelle Politikberatung zu fundieren, stimmen 76% zu. Gefragt nach Problemen und Mängeln beim Datenzugang geben 76 % die unzulängliche Bereitstellung vorhandener Daten an und 70 % fehlende Verknüpfungsmöglichkeiten. 40 % der Befragten mit Arbeitsgebiet Bildungsökonomik geben als Problem an, dass Daten in Deutschland nicht existent sind. Dieser Wert liegt deutlich über dem Durchschnitt von 29 % in anderen Arbeitsgebieten.

Am Beispiel des Schulbereichs werden diese Defizite schnell deutlich: Amtliche Statistiken liegen für den Schulbereich in den meisten Bundesländern nicht in Form von Schüler-individualdaten vor. Individuelle Bildungsverlaufsdaten gibt es nur regional vereinzelt. Insofern sie vorhanden sind, ist der Zugang der Wissenschaft zu diesen amtlichen Mikrodaten meist mit hohen Hürden verbunden. Kennzahlen zur Soziodemografie der Schülerschaft oder Erfolgsindikatoren wie Übergangsraten und Bildungsabschlüsse liegen in der Regel nur in aggregierter Form vor. Noch größer sind die Hürden, wenn mit Kompetenzdaten geforscht werden soll. Zwar wurden hier mit dem Zugang der Forschung zu den IQB-Bildungstrenddaten Fortschritte erzielt, dennoch sind vergleichende flächendeckende Kompetenzerhebungen wie z.B. die VERA-Daten der Forschung nach wie vor nur in einigen Bundesländern zugänglich. Außerdem ist es auch von politischen Entscheidungen abhängig, ob Daten für Forschungsfragen verfügbar gemacht werden. So kann die Nutzung der IQB-Bildungstrenddaten für sogenannte „neuartige“ Ländervergleiche auch von der Zustimmung der KMK abhängen.³

Diese Stellungnahme fügt sich in eine Reihe von Forderungen nach besseren Bildungsdaten und einem regelhaften Zugang der Forschung zu Bildungsdaten ein.⁴ Über bildungswissenschaftliche Disziplinen hinweg fordert die Bildungsforschung seit geraumer Zeit ein **Bildungsverlaufsregister**, in dem individuelle Bildungskarrieren **pseudonymisiert** über Bildungsabschnitte und -institutionen hinweg beobachtet werden können.⁵ Die Forderung nach einem solchen Bildungsverlaufsregister wird jedoch nicht nur in der Wissenschaft, sondern auch von der Bildungspolitik selbst als wichtig erkannt. Bereits 2003 wurde von der Kultusministerkonferenz (KMK) ein Kerndatensatz für den Schulbereich beschlossen, der die schulische Laufbahn von Schülerinnen und Schülern abbilden soll. Dieses Vorhaben ist allerdings noch nicht umgesetzt.⁶

Für die Umsetzung und Akzeptanz eines Bildungsverlaufsregisters ist es unabdingbar, datenschutzrechtliche Bedenken zu adressieren. **Datenschutz muss und kann auf höchstem Niveau gewährleistet werden.** Ein Bildungsverlaufsregister bedeutet nämlich ausdrücklich nicht, die Bildungswege von Individuen gegenüber Dritten und damit auch Akteuren der

³ Vgl. <https://www.iqb.hu-berlin.de/fdz/Datenzugang/SUF-Antrag/neuLVs/>

⁴ Vgl. z.B. BMWK (2016). Mehr Transparenz in der Bildungspolitik. <https://www.bmwk.de/Redaktion/DE/Publikationen/Ministerium/Veroeffentlichung-WissenschaftlicherBeirat/wissenschaftlicher-beirat-mehr-transparenz-in-der-bildungspolitik.pdf?blob=publicationFile&v=6>

⁵ Vgl. Rat für Sozial- und Wirtschaftsdaten (2022). Positionspapier des RatSWD: Aufbau eines Bildungsverlaufsregisters: Datenschutzkonform und benutzerfreundlich: Empfehlungen des Rates für Sozial- und Wirtschaftsdaten (RatSWD) für den Aufbau eines Bildungsverlaufsregisters. Berlin. <https://www.konsortswd.de/wp-content/uploads/Positionspapier-RatSWD-Aufbau-eines-Bildungsverlaufsregisters.pdf>

⁶ Vgl. KMK (2020). Ländervereinbarung über die gemeinsame Grundstruktur des Schulwesens und die gesamtstaatliche Verantwortung der Länder in zentralen bildungspolitischen Fragen. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2020/2020_10_15-Laendervereinbarung.pdf

öffentlichen Hand offenzulegen. Nicht der „gläserne Schüler“ ist das Ziel, sondern der Politik auf der Basis exzellenter Forschung empirische Befunde für Entscheidungen liefern zu können. Ein Rückschluss auf einzelne Individuen wird dabei durch geeignete technische und rechtliche Mittel ausgeschlossen. Dass dies möglich ist, zeigt der Blick ins europäische Ausland.

Die vorliegende Stellungnahme besteht aus vier Abschnitten:

- 1 Es wird argumentiert, dass die verfügbaren Daten unzureichend sind, um die von Bund, Ländern und Kommunen vorgesehenen Mittel, beispielsweise für Bildungsprogramme, **zielgerichtet** zu verteilen und geplante umfängliche und wissenschaftlich fundierte Wirkungsanalysen oder gar **Kosten-Effektivitäts-Analysen** umzusetzen.
- 2 Es wird beschrieben, wie die bisherige Datenbasis als Grundlage für ein **Bildungsverlaufsregister** genutzt werden könnte und wie noch bestehende Lücken entlang der verschiedenen Abschnitte und Bildungsorte eines individuellen Bildungsverlaufs zu schließen sind.
- 3 Ein Blick in das **europäische Ausland** zeigt, dass dort das geforderte Bildungsverlaufsregister schon vielerorts **datenschutzkonforme Realität** ist.
- 4 Aus den Überlegungen werden Forderungen an ein Bildungsverlaufsregister in Deutschland formuliert.

2. Fehlende Daten und Politikevaluation

Die Bildungsausgaben Deutschlands sind im OECD-Vergleich niedrig: mit 4,3 % des BIP liegen sie unter dem OECD-Durchschnitt (4,9 %). Auch wenn die öffentlichen Bildungsausgaben absolut und relativ über die Zeit gestiegen sind⁷, bleiben die Mängel des Bildungssystems offensichtlich. Die IQB-Bildungstrends 2022 zeigen erneut erhebliche Probleme insbesondere im Fach Deutsch auf: Ein Drittel der Neuntklässler erreicht den Mindeststandard im Lese- und Hörverständnis nicht.⁸ Gleichzeitig gilt das deutsche Bildungssystem als eines, in dem der Bildungserfolg stärker von der sozialen Herkunft abhängt als in anderen Ländern.

Das Problem erscheint politisch erkannt und der Wille zum Gegensteuern ist da. Ein aktuelles Beispiel für ein solches Vorhaben ist das angekündigte „Startchancen-Programm“ der Bundesregierung, das mit einem Volumen von jeweils 2 Mrd. Euro über 10 Jahre insbesondere benachteiligte Schülerinnen und Schüler unterstützen soll. Mehr Geld führt aber nicht automatisch zu besserer Bildung, denn nur ein **zielgerichteter Einsatz** kann dazu beitragen, die gewünschten Ergebnisse wie z.B. das Erreichen von Bildungsstandards oder auch der Abbau bestehender Chancenungleichheiten zu erreichen. Die Zielerreichung und der effiziente Einsatz der Mittel sollten jedoch in einer **Evaluation nach wissenschaftlichen Standards**

⁷ OECD (2022), Education at a Glance 2022: OECD Indicators, OECD Publishing, Paris, <https://doi.org/10.1787/3197152b-en>

⁸ Vgl. IQB (2023). IQB-Bildungstrend 2022. <https://www.iqb.hu-berlin.de/bt/BT2022/Bericht>

überprüft werden.⁹ Eine wissenschaftliche Begleitung und Evaluation sind als Teil des Startchancen-Programms vorgesehen, doch sowohl für die bedarfsgerechte Mittelvergabe als auch für eine Evaluation der Wirkung und der Effizienz von Bildungsprogrammen fehlt in Deutschland vielfach die notwendige Datenbasis, wie im Folgenden dargelegt wird.

Die zielgerichtete und bedarfsgerechte Mittelvergabe

Für eine zielgerichtete Mittelvergabe benötigt die Bildungspolitik Informationen über Bedarfe von Bildungseinrichtungen. Dies gilt auch für die Unterstützung benachteiligter Schülerinnen und Schüler, die das Startchancen-Programm erreichen möchte. Die Bedarfe können über verschiedene Merkmale definiert werden, darunter der soziale Hintergrund der Familien oder auch nachweisbare Defizite im Kompetenzerwerb der Bildungsteilnehmerinnen und -teilnehmer. Diese relevanten Kennzahlen liegen in der Regel nicht auf ausreichend kleinräumiger Ebene, d.h. auf der Ebene der Kindertagesstätten¹⁰, der (beruflichen) Schule, der Hochschule oder gar auf der Ebene der Bildungsteilnehmerinnen und -teilnehmer vor.

Für die Verteilung von Mitteln im Bildungsbereich wird bislang häufig der Königsteiner Schlüssel verwendet, der sich in den Flächenstaaten weitgehend an der Bevölkerungszahl orientiert und derzeit die Stadtstaaten bevorzugt.¹¹ Indikatoren für Bedarfe, wie der Anteil von Schülerinnen und Schüler ohne gute Deutschkenntnisse, der Anteil armutsgefährdeter Schülerinnen und Schüler oder die Wirtschaftskraft der Region (vor Ausgleichszahlungen), bleiben unberücksichtigt. Dadurch werden öffentliche Mittel undifferenziert verteilt – eine verbesserte Datengrundlage könnte eine zielgerichtete Mittelvergabe ermöglichen.

Auch wenn die konkrete Umsetzung des Startchancen-Programms derzeit noch zwischen Bund und Ländern verhandelt wird, steht bereits fest: Das Startchancen-Programm strebt einen gezielten Mitteleinsatz für 4.000 besonders benachteiligte Schulen an, deshalb soll erstmals nicht der Königsteiner Schlüssel, sondern Merkmale wie der Anteil Jugendlicher mit nicht deutscher Herkunftssprache, die Armutsgefährdungsquote und das Bruttoinlandsprodukt zur Bestimmung der Mittelverteilung zwischen den Ländern genutzt werden. Die Auswahl der geförderten Schulen soll durch die Länder erfolgen. Einen einheitlich definierten Sozialindex, den man als Kriterium für die Mittelverteilung nutzen könnte, gibt es in Deutschland aber nicht. Um dem Ziel der gezielten Mittelvergabe an Einzelschulen gerecht zu werden, würde aber auch ein regionaler Sozialindex bspw. auf Ebene der Kommunen nicht ausreichen. Es bedarf eines schulscharfen Sozialindex, denn Bildungsungleichheit zeigt sich insbesondere auf Ebene der Schulen innerhalb einer Kommune.

Welche Informationen/Daten wären für einen solchen schulscharfen Sozialindex erforderlich? Auf Ebene der Schülerinnen und Schüler wären dies Informationen zum elterlichen Hintergrund, einschließlich Migrations- und Bildungshintergrund sowie Sozialleistungsbezug.

⁹ Vgl. auch SWK (2022). Entwicklung von Leitlinien für das Monitoring und die Evaluation von Förderprogrammen im Bildungsbereich. Berlin.

¹⁰ Im Folgenden: Kitas

¹¹ https://www.kmk.org/fileadmin/Dateien/pdf/KMK/SWK/2022/SWK-2022-Impulspapier_Monitoring.pdf

Können die Daten nicht für die individuellen Schülerinnen und Schüler erhoben werden, so könnten über Adressinformationen der Bildungsteilnehmerinnen und -teilnehmer raumbezogene Daten der Kommunen angespielt werden, ohne dass die auswertende Forschung Zugang zu den Adressen erhält. Regelmäßige Kompetenzmessungen im Zeitverlauf würden zudem erlauben, auf Veränderungen im (relativen) Lernerfolg einzelner Schulen zu reagieren. Nur eine Kombination dieser Daten kann die Grundlage dafür schaffen, schulspezifische Bedarfe gezielt zu ermitteln.

Die Wirkungsanalyse

Soll Bildungspolitik nicht nur bedarfsgerecht, sondern auch wirkungsorientiert sein, muss der Mitteleinsatz in Hinblick auf die erlangten Kompetenzen und Bildungsbiografien der Schülerinnen und Schüler evaluiert werden. Das gilt insbesondere für Programme, die den Bildungseinrichtungen Autonomie in der Verwendung dieser Mittel geben.

Für eine Kausalanalyse der Bildungswirkungen fehlt jedoch die flächendeckende Datenbasis, insbesondere zu Kompetenzen und zu individuellen Bildungskarrieren, inklusive Bildungsübergängen und -abschlüssen. Diese liegen in der Regel nur aggregiert auf Länderebene vor. Eine über das Bildungssystem hinausgehende Analyse, etwa über langfristige Wirkungen auch auf berufliche Laufbahnen oder die Gesundheit, ist aufgrund fehlender verknüpfter Datenquellen erst recht nicht möglich.

Aufgrund fehlender Daten beschränkt sich Begleitforschung der Bildungssteuerung weitestgehend auf eine **inputbasierte Steuerung** über detaillierte Verwendungskonzepte der Bildungseinrichtungen. Ob nach dem Mittelfluss die damit finanzierten Konzepte ihre angedachte Wirkung entfalten, wird in der Regel nicht überprüft. Damit ist weder die Effektivität noch die Effizienz der Mittelvergabe sichergestellt.

Wie auch bei der bedarfsgerechten Mittelvergabe sind umfassende und kontinuierlich erhobene Daten zum elterlichen Hintergrund (einschließlich Migrations- und Bildungshintergrund sowie Sozialleistungsbezug) für faire Vergleiche zwischen Schulen erforderlich. Nur so können Schulen mit ähnlicher Ausgangslage hinsichtlich des erreichten Bildungserfolgs verglichen werden. Dazu müssen auf Ebene der einzelnen Bildungsteilnehmerinnen und -teilnehmer Bildungsverläufe erfasst werden und **wiederholte flächendeckende standardisierte Kompetenzmessungen** durchgeführt werden. Bildungserfolg manifestiert sich allerdings nicht nur über Kompetenzmaße im engeren Sinne und Bildungsverläufe. Für eine Evaluation der Wirkung der Mittel oder eines Programms auf den langfristigen Bildungserfolg ist eine Verknüpfung auch etwa mit Daten der Arbeitsmarktstatistik erforderlich. Liegen diese Daten vor, ist es mit geeigneten statistischen Verfahren möglich, den kausalen Zusammenhang zwischen Mittelvergabe und Bildungserfolg herauszuarbeiten.

Kosten-Effektivitäts-Analysen

Selbst wenn die Evaluation von Bildungsprogrammen deren Wirksamkeit nachweisen kann, sollten die Effekte verschiedener Maßnahmen miteinander verglichen werden, um öffentliche Gelder dort einzusetzen, wo sie die **größtmögliche Wirkung** entfalten. Um dies zu erreichen, müssen neben der Wirkungsanalyse auch Informationen zu den umgesetzten Maßnahmen und deren Kosten erfasst werden. Eine Kombination der Verwendungskonzepte, der entstehenden Kosten und der hier skizzierten Wirkungsanalyse würde eine Kosten-Effektivitäts-Analyse ermöglichen – sowohl innerhalb als auch zwischen Programmen.

3. Datenbasis für ein Bildungsverlaufsregister

Der vorherige Abschnitt hat aufgezeigt, wie wichtig gute Daten für eine evidenzbasierte und effiziente Bildungspolitik sind. Im Folgenden wird die hierfür notwendige Datenbasis im Detail beschrieben. Zudem wird überprüft, auf Basis welcher bereits bestehenden Daten ein Bildungsverlaufsregister erstellt werden könnte und wie es um weitere Datenquellen ergänzt werden kann. Der anschließende Blick ins europäische Ausland zeigt, dass ein solches Bildungsverlaufsregister bereits gelebte datenschutzkonforme Realität in vielen europäischen Ländern ist.

Ein Bildungsverlaufsregister: Datenschutzkonform über alle Bildungsabschnitte

Um Mittel effektiv und effizient für die Weiterentwicklung des Bildungssystems einsetzen zu können, benötigen Bildungsforschung, -politik und -administration Zugriff auf umfassende Längsschnittdaten über individuelle Bildungsverläufe – ein Bildungsverlaufsregister.¹² In einem Bildungsverlaufsregister werden Daten aus verschiedenen Bildungsstatistiken mithilfe einer **pseudonymisierten Identifikationsnummer** zusammengeführt. Auf diese Weise werden Bildungsverläufe verknüpft, um Analysen über verschiedene Bildungsbereiche hinweg zu ermöglichen. Ein Bildungsverlaufsregister ist dabei kein einzelner Datensatz, in dem alle Bildungsabschnitte und Merkmale verknüpft sind. Vielmehr können die Einzeldatensätze dezentral (über Bildungsabschnitte, datenhaltende Gebietskörperschaften) vorgehalten werden. Wichtig ist jedoch, dass diese Einzeldatensätze projekt- und fragestellungsbezogen verknüpft werden können.

Eine Zusammenführung der Datensätze erfolgt durch einen **Datentreuhänder** und nur in dem Maße, wie es zur Beantwortung einer Fragestellung notwendig ist. Auch werden die Daten nur in anonymisierter Form über sichere Zugangswege an Forschende weitergegeben. Ergebnisse der Forschung werden in aggregierter Form publiziert und ein Rückschluss auf einzelne Individuen wird durch geeignete technische und rechtliche Mittel ausgeschlossen. Die

¹² Vgl. RatSWD, 2022; Giar, K, Hohlstein, F., Wipke, M., Scharnagl, A., (2023). Konzeption eines Statistischen Bildungsverlaufsregisters in Deutschland – Entwicklungen bis 2023 und Gestaltungsoptionen. Statistisches Bundesamt, WISTA 3. 51-62.

Bereitstellung der Daten wird mit Kosten verbunden sein und die zu schaffende Funktion des Datentreuhänders muss daher mit ausreichenden finanziellen Ressourcen ausgestattet sein. Hier könnte die Rolle des Forschungsdatenzentrums der Statistischen Ämter der Länder (FDZ Länder) ausgebaut werden.

Wichtig ist, dass das Bildungsverlaufsregister **sämtliche Bildungsabschnitte** umfasst. Bildung beginnt mit der frühkindlichen Bildung. Diese findet zumeist in der **Familie** und in der **Kindertagesbetreuung** statt. Der frühkindlichen Bildung schließt sich die formale Schullaufbahn über **Grundschule und Sekundarschule** an, doch ein Bildungsverlaufsregister muss auch über die grundlegende Schullaufbahn hinaus daran anschließende Bildungswege über **Berufsbildung** und **Hochschullaufbahn** abbilden. Eine Möglichkeit der Erfassung **lebenslangen Lernens über betriebliche oder außerbetriebliche Weiterbildung** ist ebenfalls wünschenswert.

Hinsichtlich des Merkmalkranzes sollten keine unnötig engen Grenzen gesetzt werden. Dies ist wichtig, da aus heutiger Sicht nicht relevante Merkmale künftig eine wichtige Rolle in Forschungsprojekten spielen können. Eine registerbasierte Lösung sollte stattdessen Merkmale, die im Laufe der administrativen Prozesse erhoben werden, erfassen und mit entsprechenden Auswertungs- und Forschungskonzepten der Wissenschaft zur Verfügung stellen.¹³ Wichtig ist, dass **soziodemografische Merkmale** der Bildungsteilnehmerinnen und -teilnehmer und im Falle von Minderjährigen auch die der Eltern erfasst werden. Dazu gehören Geschlecht, Alter, Zuwanderungsgeschichte sowie Sozialleistungsbezug und Bildungsniveau der Eltern. Der individuelle **Bildungserfolg** muss über Übergänge (auf Schulen, Berufsschulen und Hochschulen), erlangte Abschlüsse, Studienerfolg bzw. -abbruch sowie über weitere Leistungsmaße erfasst sein. Insbesondere der Erfassung von **Kompetenzen** kommt eine wichtige Bedeutung zu.

Für viele Fragestellungen von hoher gesellschaftlicher Relevanz ist zudem eine Verknüpfbarkeit der Daten über das Bildungsverlaufsregister hinaus erforderlich. Bildungserfolg misst sich nicht nur in Kompetenzmaßen und Bildungsabschlüssen. Eine gute Bildungspolitik kann Arbeitsmarkt-, Wirtschafts-, Sozial-, Gesundheits-, Integrations- und Familienpolitik zugleich sein, und bessere Bildung und geringere Bildungsungleichheit können Auswirkungen auf Armut, Gesundheit, Kriminalität und vieles mehr haben. Die Bestimmung des gesellschaftlichen Mehrwerts besserer Bildung verlangt nach einer **Verknüpfung mit Datenregistern anderer Gesellschaftsbereiche**, etwa mit Arbeitsmarkt- und Gesundheitsdaten. Ebenso wünschenswert ist die Möglichkeit einer **Verknüpfung mit weiteren umfragebasierten Datensätzen**, um Lücken der Registerdaten gezielt für spezielle Forschungsfragen schließen zu können.

¹³ Registerbasierte Daten sind Informationen, die aus offiziellen staatlichen oder öffentlichen Registern und Verwaltungsdokumenten stammen und soziale oder demografische Merkmale wie Bevölkerung, Arbeitsmarkt, Bildung und Gesundheit dokumentieren. Sie werden von staatlichen Stellen erfasst und dienen als zuverlässige Grundlage für Forschung, Statistik und politische Entscheidungsfindung.

Bildungs(verlaufs)daten: Eine Bestandsaufnahme

Generell werden in Deutschland bildungsbezogene Daten oft nur in Stichproben und als Querschnitte erhoben. Sie liegen häufig in **nicht verknüpfbaren Datensilos** und sind oft nur nach dem Überwinden **hoher bürokratischer und administrativer Hürden** für die Forschung nutzbar. Der folgende Überblick entlang der einzelnen Bildungsabschnitte zeigt, aus welchen Quellen ein zukünftiges Bildungsverlaufsregister gespeist werden könnte und wo neue Daten zu erheben wären.

Die Analyse von Bildungsverläufen im Zeitverlauf, auch über Bildungsabschnitte und Institutionen hinweg, ist bisher nur durch **umfrage- und kohortenbasierte Datensätze**, wie zum Beispiel das Nationale Bildungspanel (NEPS), möglich. Das NEPS stellt dabei eine Reihe von Längsschnittdatensätzen verschiedener Kohorten zur Verfügung. Eine Besonderheit der NEPS-Daten ist die Möglichkeit, einige Individualdatensätze mit den Daten der Sozialversicherung des IAB zu verknüpfen, die wiederum detaillierte registerbasierte Informationen zu individuellen Erwerbsverläufen und Sozialleistungsbezug beinhalten. Diese Verknüpfung ermöglicht somit die Messung langfristiger Bildungseffekte über die berufliche Bildung hinaus und zeigt so die Potentiale für die Bildungsforschung und -steuerung, die durch die Verknüpfbarkeit von Bildungs- und Arbeitsmarktdaten entstehen. Auch bietet das NEPS, ähnlich wie andere repräsentative Befragungen des SOEP, FReDa und AID:A, Informationen zu den bildungsrelevanten Aktivitäten einer Familie. Durch eine Verknüpfung derartiger Daten wäre es möglich, den Bildungsort Familie in Analysen zum Bildungsverlauf von Kindern und Jugendlichen zu berücksichtigen.

Bildungsort Kindertagesbetreuung

Daten zur frühen Bildung in der **öffentlich finanzierten Kindertagesbetreuung**¹⁴ umfassen Daten zu dem Bereich der Kindertageseinrichtungen der öffentlich finanzierten Kindertagespflege und auch die Ausgaben und Einnahmen der Kinder- und Jugendhilfe für diese Angebote. Die Erhebungen haben ihre gesetzliche Verankerung im SGB VIII in den §§ 98 – 103 und erfassen Angaben über das Personal in den Einrichtungen, die Anzahl und Alter der betreuten Kinder, Migrationshintergrund, die zu Hause mehrheitliche gesprochene Sprache und auch den Anteil der Kinder mit Förderbedarf. Angaben zu den Kosten einzelner Bildungsanbieter liegen nicht vor.

Wie auch in anderen Bildungsabschnitten beinhalten die amtlichen Daten keine Erfassung von Entwicklungsmaßen, allerdings gibt es Datenquellen, die sich bei flächendeckender Umsetzung der Erhebungen eignen würden, die amtlichen Daten zu ergänzen. Eine mögliche Datenquelle zur frühkindlichen Entwicklung sind in vielen Bundesländern Sprachstanderhebungen, die punktuell in Kindertageseinrichtungen stattfinden. Erhebungsmethoden, Erhebungsrhythmen

¹⁴ Vgl. C. Katharina Spieß (2009). Improvements and Future Challenges for the Research Infrastructure in the Field of "Preschool Education". Working Paper of the Council for Social and Economic Data (Rat für Sozial- und Wirtschaftsdaten) Nummer 60. Berlin. https://www.ratswd.de/download/workingpapers2009/60_09.pdf

und die Datenverfügbarkeit für die Forschung sind jedoch sehr unterschiedlich in den verschiedenen Ländern geregelt.¹⁵

Eine weitere mögliche Datenquelle zur Erfassung frühkindlicher Kompetenzen stellen flächendeckende Schuleingangsuntersuchungen am Übergang von vorschulischer zur schulischen Bildung dar. Aber auch hier ist der Datenzugang bisher weitgehend intransparent und über die Bundesländer uneinheitlich geklärt.¹⁶ Weitergehende Informationen zum vorschulischen Entwicklungsstand lassen sich aus Befragungsdaten ableiten.

Bildungsort Schule

Eine **Schülerindividualstatistik** für allgemeinbildende sowie berufsbildende Schulen gibt es bislang nicht in allen Bundesländern. Zudem ist ein Zugang zu den Individualdaten, die zumeist nur im Querschnitt vorliegen und einen begrenzten Merkmalskranz haben, für die Forschung mit hohen Hürden verbunden. Die Statistik der allgemeinbildenden Schulen erfasst in regionalen Aggregaten die Anzahl von Schulen, Klassen, Schülerinnen und Schüler, Absolvierenden und Abgehenden sowie deren Lehrkräfte. Daten zu Kompetenzen sind nicht enthalten.

Insbesondere in einem mehrgliedrigen Schulsystem kommt Übergängen zwischen Bildungsabschnitten eine besondere Bedeutung zu. Übergangsraten zwischen Bildungsabschnitten und Schulformen werden jedoch **nur in aggregierter Form** von den Statistischen Landesämtern zur Verfügung gestellt. Für einzelne Grundschulen ist so ersichtlich, wie viele Kinder auf welche weiterführenden Schulformen wechseln. So beinhalten sie – wenn auch nur indirekt – Informationen darüber, welche Schulformen in welchen Nachbarschaften überwiegend besucht werden und welche Bildungsabschlüsse dort dominieren. Es ist jedoch mit diesen Daten nicht möglich, Bildungsverläufe zu analysieren. Dies würde erst durch den von der KMK beschlossenen **Kerndatensatz** möglich werden.

Eine große Lücke der bestehenden amtlichen Statistik besteht durch den Mangel an Daten zu flächendeckenden **Kompetenzmessungen**. Um auf Kompetenzmessungen zurückgreifen zu können, nutzt die Bildungsforschung und -politik in Deutschland vor allem internationale Vergleichsstudien und nationale Stichproben, die aber nicht im Längsschnitt erhoben werden. Die Daten aus Schulleistungsstudien wie z. B. PISA, IGLU oder die IQB-Ländervergleiche werden der Wissenschaft für Re- und Sekundäranalysen zur Verfügung gestellt. Dabei unterliegen Forschungsanträge, die Vergleiche zwischen den Bundesländern anstreben, einer besonderen Begründung und können auch abgelehnt werden. Für eine effektive Steuerung und als wichtiges Element für das Bildungsverlaufsregister wären allerdings standardisierte,

¹⁵ Vgl. Autorengruppe Bildungsberichterstattung (2020). Bildung in Deutschland 2020. S. 4. Verfügbar unter: www.bildungsbericht.de

¹⁶ Vgl. Hübener, M., Kühnle, D., Spieß, C.K. (2019). Die Verwendung von Schuleingangsdaten für bildungs- und familienökonomische Analysen – das Beispiel der Analyse von Elterngeldeffekten auf sozioökonomische Unterschiede bei kindlichen Entwicklungsmaßen, in: Maik-Carsten Begemann und Klaus Birkelbach (Hrsg.): Forschungsdaten für die Kinder- und Jugendhilfe, Qualitative und quantitative Sekundäranalysen, Springer Verlag, Wiesbaden, S. 457-474.

regelmäßige und flächendeckende Kompetenzmessungen unerlässlich. Dazu könnten die bereits bestehenden aber für die Forschung bisher kaum nutzbaren **Vergleichsarbeiten** (VERA) unter Nutzung der stichprobenbasierten Erhebungen der IQB-Bildungstrends und NEPS weiterentwickelt werden. Durch eine Verknüpfung von flächendeckenden Kompetenzmessungen mit den Individualdaten eines Kerndatensatzes ließe sich eine der wichtigsten Datenlücke der amtlichen Schulstatistik ohne großen Mehraufwand schließen.

Bildungsort Hochschule

Amtliche Gesamterhebungen im Hochschulbereich liegen über die Statistik der Studierenden und die Statistik der Prüfungen vor. Diese Bundesstatistiken erfassen die Grundgesamtheit aller Studierenden mit Fach, Dauer und Institution sowie aller Abschlussprüfungen. Beide Statistiken liegen bislang nur im Querschnitt vor und der Merkmalskranz ist sehr reduziert. Die Daten enthalten keine Informationen zur vorangegangenen Bildung wie z.B. der Note der Hochschulzugangsberechtigung oder zum Schulabschluss oder auch vorangegangenen Berufsausbildungen. Auch Leistungsdaten sind mit Ausnahme der Abschlussnote nicht verfügbar. Mit der Novelle des Hochschulstatistikgesetzes von 2016 wurde erstmals eine **Studienverlaufsstatistik** eingeführt, die es in Zukunft ermöglicht, individuelle Studienverläufe auch über verschiedene Institutionen und studierte Fächer zu verfolgen. Diese Statistik steht der Wissenschaft aktuell noch nicht für Auswertungen zur Verfügung, sie sollte aber in das Bildungsverlaufsregister integriert werden. Dringenden Anpassungsbedarf gibt es bei den Löschungsvorschriften, denn die Pseudonyme und die Zusammenführungen müssen aktuell nach 18 Jahren gelöscht werden, nach einer Promotion sogar bereits nach vier Jahren. Aus Sicht der Wissenschaft sind solche **Löschungsvorschriften** ein Problem. Hier wird enormes Analysepotential verschenkt und eine Anpassung der gesetzlichen Regelungen im Hochschulstatistikgesetz ist erforderlich.

Bisher nicht abgebildet in der amtlichen Statistik ist der Übergang zwischen Schule und Hochschule. Dieser könnte im Bildungsverlaufsregister durch Daten der Plattform "Hochschulstart" ergänzt werden.¹⁷ Zwar nehmen nicht alle Hochschulen am Verfahren zur **Studienplatzvergabe** durch Hochschulstart teil, ein großer Teil der Hochschulen wäre jedoch abgedeckt und könnte Informationen zum Thema Präferenzen und Studienplatzvergabe beitragen. Die Daten sollten, auch wenn sie nicht Teil des Bildungsverlaufsregisters werden, als zu verknüpfende Daten und regelhaft für die Forschung zur Verfügung stehen.

Bildungsort Betrieb und Berufsschule

Einen wichtigen Stellenwert im deutschen Bildungssystem hat die **Berufsbildung im dualen System**. Sie muss dementsprechend ein Teil eines Bildungsverlaufsregisters sein. Auch hier bietet die amtliche Statistik mit der Statistik der beruflichen Schulen aktuell nur aggregierte Querschnittsdaten. Der betriebliche Teil der Ausbildung wird über die **Berufsbildungsstatistik** abgedeckt, die die Grundgesamtheit aller in Deutschland erfassten Ausbildungsverträge (nicht: Auszubildende) umfasst und Angaben über Art der Ausbildung, Dauer und Erfolg enthält.

¹⁷ <https://hochschulstart.de/startseite/>

Weitere stichproben- und umfragebasierte Datensätze zur betrieblichen Ausbildung stehen über das Bundesinstitut für Berufsbildung (BIBB) zur Verfügung und sollten perspektivisch in das Register integrierbar sein.

Lebenslanges Lernen

Bildungsinvestitionen außerhalb der sekundären und tertiären Bildungssysteme sind in Deutschland zunächst auf Anbieterseite erfasst und bieten nur wenige Informationen, die einem individuellen Bildungsverlaufsregister zugespielt werden könnten (etwa die Anbieterstatistiken und die Statistik der Volkshochschulen des Deutschen Instituts für Erwachsenenbildung). Eine administrative Datenquelle auf Teilnehmerseite, die sich für eine Einbindung in ein Bildungsverlaufsregister anbieten würde, liegt jedoch nicht vor.

Verknüpfbarkeit der Daten

Die bisherigen Ausführungen haben gezeigt, dass bereits für die meisten Bildungsstufen Daten der amtlichen Statistik oder anderer öffentlicher Anbieter vorliegen, die sich als Module für den Aufbau eines Bildungsverlaufsregistern anbieten oder zu solchen weiterentwickelt werden könnten. Diese Datensätze sind dezentral organisiert, eine Verknüpfung dieser voneinander getrennten Datensätze durch einen Datentreuhänder ist aber umsetzbar. Dazu muss eine **einheitliche Personen-ID** eingeführt werden. Das ebenfalls diskutierte *Record-linkage* über unveränderliche Personendaten ist fehleranfällig und deshalb keine langfristig zielführende Alternative. Wenn die für den Registerzensus der amtlichen Statistik anvisierte Personen-ID, nämlich die Steuer-ID, nicht auch im Bildungsverlaufsregister genutzt werden kann, muss bei einer bereichsspezifischen Personen-ID unbedingt die perspektivische Verknüpfbarkeit mit anderen Registern oder stichprobenbasierten Erhebungen mitgedacht werden.

4. Keine Utopie, sondern europäische Realität

Ein Bildungsverlaufsregister ist bereits **datenschutzkonforme Realität in vielen europäischen Ländern**, von denen einige im Folgenden exemplarisch hinsichtlich der Bereitstellung von Bildungsverlaufsdaten beschrieben werden.

So stellen u.a. die skandinavischen Länder, die Niederlande, Österreich, die Schweiz und das Vereinigte Königreich umfassende und flächendeckende Bildungsverlaufsdaten der Bildungsforschung und -steuerung zur Verfügung¹⁸ und ermöglichen dort eine evidenzbasierte und wirkungsorientierte Bildungspolitik. Erfahrungen aus und Beispiele zur institutionellen Ausgestaltung der Dateninfrastruktur in diesen Ländern können helfen, einen Aufbau einer entsprechenden Bildungsdateninfrastruktur auch in Deutschland zu ermöglichen bzw. zu beschleunigen.

¹⁸ Weitere Informationen gibt es z.B. auf der Übersichtseite zum Bildungsverlaufsregister im Vereinigten Königreich: <https://find-npd-data.education.gov.uk/categories> oder zu den amtlichen Mikrodaten (inkl. der Bildungsverlaufsdaten) der Niederlande: <https://www.cbs.nl/en-gb/our-services/customised-services-microdata/microdata-conducting-your-own-research/microdata-catalogue>

Aus dem Vergleich der Infrastruktur europäischer Nachbarländer lassen sich Vorschläge für die institutionelle Ausgestaltung der entsprechenden Infrastruktur in Deutschland ableiten, etwa über die Verantwortung für Datenhaltung und -bereitstellung, den Umfang der Daten, deren Verknüpfbarkeit mit weiteren Datenquellen und über die Bedeutung von Kompetenzmessungen. Aus den Beispielen wird aber auch deutlich, dass die in dieser Stellungnahme formulierten Forderungen für Bildungsdaten in Deutschland nicht über in Europa übliche und längst **etablierte Standards** hinausgehen. Die Zögerlichkeit, mit der die Frage in Deutschland angegangen wird, ist aus Sicht der Wissenschaft nicht nachvollziehbar.

Datenhaltung und anonymisierte Bereitstellung

In anderen Ländern liegt die Verantwortung für Datenhaltung, Anonymisierung und datenschutzrechtlich konforme Bereitstellung zumeist bei den jeweiligen statistischen Ämtern auf nationaler Ebene. Reformen der jeweiligen nationalen Statistik-Gesetzgebung ermöglichen die registerbasierte Identifikation von Individuen durch entsprechende Identifikationsnummern über Kitas, Schulen und Bildungsstufen hinweg und erlauben so die Darstellung von Bildungsverläufen. Es gilt weitestgehend ein **Single-Entry-Prinzip**. Einmal im Register erhobene Daten stehen zur Verfügung und müssen bei zukünftigen Verwaltungsvorgängen nicht wiederholt eingetragen werden.

Die datenhaltenden Institutionen verantworten die Anonymisierung der Daten vor einer Nutzung durch Verwaltung oder Forschung für Zwecke der Bildungssteuerung oder des wissenschaftlichen Erkenntnisgewinns. Diese Anonymisierung verhindert die Nutzung der Bildungsverläufe für Zwecke, die nicht mit dem Datenschutz konform sind, etwa für illegale Profilbildung. Nach der Anonymisierung stellen die Statistikämter die Mikrodaten per Remote Access der Wissenschaft zur Verfügung.

Umfang von Bildungsdaten

Der Umfang der erfassten Daten im europäischen Ausland unterscheidet sich in drei wesentlichen Punkten von dem in Deutschland: den erfassten Bildungsstufen, der Verknüpfung mit anderen register- oder umfragebasierten Daten und dem Vorhandensein umfassender Kompetenzmessungen.

In Bezug auf die erfassten Bildungsstufen zeigen die vollständig registerbasierten Lösungen in Dänemark, den Niederlanden sowie teilweise in Österreich, der Schweiz und dem Vereinigten Königreich nur geringe Unterschiede. Die Verfolgung individueller Bildungskarrieren reicht teilweise vom vorschulischen Bereich bis zum Hochschulbereich und ermöglicht eine umfassende Darstellung von Bildungsübergängen.

Zudem ermöglichen diese umfassenden Registerlösungen die Verknüpfung mit Daten aus anderen Registern, um beispielsweise die sozioökonomische Zusammensetzung von

Schülerschaften und deren Elternhäusern zu bestimmen. Dadurch können Schulen systematisch miteinander verglichen werden und Mittelzuweisungen zielgerichtet erfolgen.

Auch Wirkungsanalysen sind in diesen registerbasierten Lösungen durch die Verfolgung von Bildungskarrieren über das Bildungssystem im engeren Sinne hinaus möglich. Somit kann nicht nur die Wirkung zusätzlicher Ressourcen auf Kompetenzen und Schulabschlüsse erfasst werden, sondern auch deren Auswirkung etwa auf den Arbeitsmarkterfolg oder die Gesundheit, da auf individueller Ebene Gesundheitsoder Arbeitsmarktdaten mit Bildungsverläufen verknüpft werden können.

Ein entscheidender Unterschied zwischen den betrachteten Ländern liegt in der Durchführung regelmäßiger und flächendeckender Kompetenzmessungen. Diese Messungen sind von großer Bedeutung für eine Vielzahl von Fragestellungen, da Bildungsübergänge, Bildungsabschlüsse und Noten allein nur begrenzt Aussagen zu Kompetenzen erlauben. Sowohl Dänemark als auch das Vereinigte Königreich führen regelmäßig nationale Kompetenzmessungen durch und stellen diese Daten verknüpft mit populationsweiten Umfrage- bzw. Registerdaten der Forschung zur Verfügung.

Aber auch im europäischen Ausland sind die Entwicklungen der Datenbestände noch nicht abgeschlossen. Um die Bedeutung von Lehrkräften für Lernerfolge zu erfassen, wird oft die Verknüpfbarkeit von Daten von Bildungsteilnehmenden und Lehrpersonal gefordert. In den meisten betrachteten Ländern war dies bisher nicht möglich. Eine Ausnahme ist Dänemark, wo individuelle Lehrkräfte in den Bildungsregistern identifiziert werden können.

Verknüpfbarkeit von Register- und Umfragedaten

Obwohl registerbasierte Systeme viele Vorteile bieten, haben sie auch ihre Grenzen, da nur solche Merkmale Eingang in die Register finden, die in Verwaltungsvorgängen erhoben werden. Diese sind nicht notwendigerweise ausreichend, um Fragestellungen der Bildungsforschung und -steuerung in vollem Umfang zu beantworten. Um diese Einschränkungen zu überwinden, besteht in einigen Ländern die Möglichkeit, die registerbasierten Bildungsdaten mit Umfragedaten zu verknüpfen. Dänemark, die Niederlande, die Schweiz und das Vereinigte Königreich nutzen bereits diese Möglichkeit und können so die Registerdaten entsprechend anreichern.

Zusammenfassend zeigt der Blick auf die verschiedenen nationalen Systeme, dass eine registerbasierte Datenbereitstellung in vielen europäischen Staaten bereits langjährige Praxis ist. Technisch und organisatorisch ist der Aufbau einer solchen datenschutzkonformen Dateninfrastruktur realisierbar. Beim Aufbau einer umfangreichen, registerbasierten Bildungsdateninfrastruktur kann Deutschland durch die erfolgreichen Modelle anderer europäischer Länder lernen.

5. Konkrete Schritte und Umsetzungsempfehlungen

Bildungspolitik braucht Forschung und Forschung braucht exzellente Daten. Die Bedeutung einer umfassenden Datenbasis für eine **effektive und effiziente Bildungspolitik** wurde in Deutschland lange wenig beachtet. In den letzten Jahren wurden zwar durch Befragungen wie dem NEPS oder den IQB-Bildungstrends Fortschritte gemacht, nach wie vor fehlt jedoch der Aufbau eines Bildungsverlaufsregisters. Wie die Ausführungen dieser Stellungnahme zeigen, sind viele Datenquellen, aus denen sich ein Bildungsverlaufsregister speisen könnte, bereits vorhanden oder es existieren konkrete Umsetzungspläne, wie beim Kerndatensatz der KMK. Die Historie des Kerndatensatzes der KMK seit 2003 zeigt aber auch, wie mühsam ein solcher Prozess in einem föderalen System sein kann. Neben den technischen Herausforderungen für das oben skizzierte Modell sind in Deutschland folglich noch große politische und juristische Hürden zu nehmen.

Vieles deutet jedoch aktuell darauf hin, dass sich die Umsetzung des Bildungsverlaufsregisters in Deutschland konkretisiert und der politische Wille (mehrheitlich) vorhanden ist. Deshalb sollte nun zeitnah mit dessen Konzeptionierung und Umsetzung begonnen werden. In anderen europäischen Ländern ist das, was in dieser Stellungnahme gefordert wird, längst Realität.

Die folgenden Empfehlungen sind dabei aus Sicht der Wissenschaft relevant:

Umfang und Abdeckung der Bildungsverläufe

Aus Sicht der Wissenschaft ist es wichtig, das Bildungsverlaufsregister nicht zu eng zu definieren, um Analysepotentiale nicht unnötig zu vergeben. Um datenschutzrechtlichen Einwänden Rechnung zu tragen, sollten Individualdaten zu einzelnen Bildungsabschnitten in separaten abschnittsbezogenen Registern gespeichert werden. Außerdem sollten alle Bildungsbereiche von der frühkindlichen Bildung bis mindestens zum Berufsbildungs- und/oder Hochschulabschluss Eingang in ein Register finden, um das gesamte Spektrum von Bildungsverläufen abbilden zu können.

Transparenter Datenzugang für die Wissenschaft

Ohne unabhängige, wissenschaftlich fundierte Analysen komplexer Bildungsverläufe durch die freie Wissenschaft fehlen der Bildungspolitik wichtige Entscheidungsgrundlagen. Der Zugang zu Daten sollte für die unabhängige Forschung klar, transparent und diskriminierungsfrei geregelt werden. Anfallende Nutzungsgebühren dürfen nicht dazu führen, dass Forschende, insbesondere Nachwuchswissenschaftlerinnen und -wissenschaftler, ihre Forschung nicht durchführen können. Ein entsprechender Datenzugang könnte z.B. über das geplante Forschungsdatengesetz geregelt werden.¹⁹

¹⁹ Vgl. Rat für Sozial- und Wirtschaftsdaten (2023). Positionspapier von RatSWD: Forschungsdatengesetz: Was zentral ist. Berlin. https://www.konsortswd.de/wp-content/uploads/RatSWD_Positionspapier_Forschungsdatengesetz_Was_zentral_ist.pdf

Verknüpfbarkeit

Die reiche Datenlandschaft in Deutschland leidet darunter, dass einzelne Datensätze unverbundene Datensilos bleiben. Das trifft auch auf Bildungsdaten zu. Ein Bildungsverlaufsregister würde bislang unverbundene Bildungsdaten verschiedener Quellen verknüpfbar machen. Dies erfordert die Einführung von Personen-IDs und einen Datentreuhänder, der die für die jeweilige Fragestellung benötigten Daten verknüpfen und für Forschungszwecke verfügbar machen kann. Auch das Bildungsverlaufsregister darf nicht selbst zum Datensilo werden. Die Analyse langfristiger Effekte von Bildung kann erst durch die Verknüpfung von Daten aus dem Bildungsverlaufsregister mit bspw. Erwerbsverläufen, Steuerdaten oder Gesundheitsdaten erfolgen. Auch die Verknüpfung mit etablierten Befragungen sollte perspektivisch ermöglicht werden. Um die Daten für die Wissenschaft bereitstellen zu können, muss die zu schaffende Funktion des Datentreuhänders mit ausreichenden finanziellen Ressourcen ausgestattet sein. Hier könnte die Rolle des Forschungsdatenzentrums der Statistischen Ämter der Länder (FDZ Länder) ausgebaut werden.

Datenschutz

Den Anforderungen der DSGVO sowie des Rechts auf informationelle Selbstbestimmung (Art. 2 Abs. 1 GG) muss und kann ohne Einschränkungen entsprochen werden. Beispiele zur Einhaltung des Datenschutzes bei Bildungsverlaufsregistern finden sich in europäischen Nachbarländern. Bei der Ausgestaltung des Zugangs zum Bildungsverlaufsregister für die unabhängige Wissenschaft muss das Forschungsprivileg für den Zugang der Wissenschaft zu sensiblen Daten Berücksichtigung finden. Zudem muss es in einem Bildungsverlaufsregister eine einheitliche Auslegung der datenschutzrechtlichen Bestimmungen geben, denn es sind sowohl Landes- als auch Bundesstatistiken betroffen. Außerdem darf es keine Löschungspflicht für Daten geben, wenn nicht sichergestellt ist, dass pseudonymisierte Daten für Bildungsverläufe verfügbar sind.

Impressum

Kontakt:

Rat für Sozial- und Wirtschaftsdaten (RatSWD)

Geschäftsstelle

Am Friedrichshain 22

10407 Berlin

office@ratswd.de

<https://www.ratswd.de>

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

Berlin, Dezember 2023



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

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

DOI: [10.17620/02671.87](https://doi.org/10.17620/02671.87)

Zitationsvorschlag:

Hertweck, F., Isphording, I. E., Matthewes, S. H., Schneider, K. & Spieß, C. K. (2023). *Für besseren Datenzugang im Bereich Bildung. Stellungnahme aus dem Verein für Socialpolitik.* (RatSWD Working Paper 282/2023). Berlin. Rat für Sozial- und Wirtschaftsdaten (RatSWD).

<https://doi.org/10.17620/02671.87>