INSTITUT FÜR ARBEITSMARKT- UND BERUFSFORSCHUNG
Die Forschungseinrichtung der Bundesagentur für Arbeit

Online Magazin IAB-Forum

20 Jahre Integrierte Erwerbsbiografien: „Für uns Forschende ist die IEB wie ein Grundnahrungsmittel“

22. Dezember 2023 | Christiane Keitel



Das IAB ist nicht nur ein Hort der Forschung, sondern auch einer der größten Produzenten von Daten zur Arbeitsmarktforschung, die weltweit von Forschenden genutzt werden. Einen der bekanntesten Datensätze des Instituts, die Integrierten Erwerbsbiografien (IEB), gibt es nun seit 20 Jahren. Anlässlich dieses Jubiläums hat die Redaktion des IAB-Forums bei Ali Athmani und Andreas Schneider nachgefragt, die den Datensatz seit Jahren aufbereiten und bereitstellen, aber auch bei Gesine Stephan, die als Forscherin und Nutzerin die IEB genau kennt.

Liebe Kollegen aus dem Daten- und IT-Management, Sie kennen den Datensatz wohl mit am besten: Was ist das Besondere an den Integrierten Erwerbsbiografien?

Ali Athmani: Ich kenne keinen anderen Datensatz in Deutschland, der so viele unterschiedliche Mikroinformationen zum Arbeitsmarkt vereint. Die IEB speist sich aus Datenquellen zu Beschäftigung, Leistungsbezug, Arbeitsuche und Maßnahmeteilnahme und gibt damit Einblick in unterschiedlichste Aspekte. Aber nicht nur die Vielfalt, sondern auch die zeitliche Dimension der Daten ist einzigartig: Dadurch, dass die IEB von den 1970er Jahren bis heute reicht, kann man individuelle Erwerbsverläufe über lange Zeit verfolgen. Und die Forschenden können damit – im Unterschied zu Umfragen – auch kleine Teilgruppen des Arbeitsmarktes mit guter Qualität empirisch untersuchen. Oder im Vergleich zu stichtagsbezogenen Daten sehr flexible, biografisch orientierte Auswertungen durchführen. Wir haben hier also einen unglaublich umfangreichen Datensatz, der eine solide Grundlage für die evidenzbasierte Forschung darstellt.



Prof. Dr. Gesine Stephan (Leiterin des Forschungsbereichs „Arbeitsförderung und Erwerbstätigkeit“), Ali Athmani (Leiter des Geschäftsbereichs Daten- und IT-Management) und Andreas Schneider (Mitarbeiter im Daten- und IT-Management) (von links nach rechts).

Wie entsteht so ein Datensatz genau?

Andreas Schneider: Die IEB steht am Ende einer Kette von Datenprüfungen und -transformationen: Die eigentlichen Datenproduzenten sind ja das Fachpersonal in Betrieben, Arbeitsagenturen und Jobcentern. Deren Eingaben erhält zunächst das DataWarehouse der Bundesagentur. Dort strukturieren die Kolleginnen und Kollegen die Daten und legen sie so ab, dass sie für verschiedenste Auswertungen nutzbar sind. Im IAB erzeugen wir daraus zunächst thematisch getrennte Datensätze, also etwa einen Datensatz mit Beschäftigtendaten, einen mit Arbeitsuchendendaten und so weiter. Das machen wir auch deshalb, weil die Qualität in getrennten Datensätzen forschungsspezifisch noch besser optimiert werden kann. Zum Schluss fügen wir die Einzelinformationen nach Zeit und Person zusammen und bereinigen bei bestimmten Merkmalen, wie zum Beispiel den Regionalmerkmalen, quellen- und zeitraumbezogene Widersprüche. Mir ist es noch wichtig zu sagen, dass der Datenschutz bei uns oberste Priorität hat. Alle Daten sind pseudonymisiert, es sind also an keiner Stelle Rückschlüsse auf Namen, Adressen oder ähnliche persönliche Informationen möglich.

> Seit es die IEB gibt, wurden über 2.000 projektspezifische Datenauszüge für Forschende erstellt.

Herr Schneider, aktuell sind Sie der Hauptansprechpartner für den Datensatz im IAB, das heißt, wer dazu Fragen oder Anregungen hat, wendet sich an Sie. Wie viele Forschende nutzen die IEB?

Schneider: Insgesamt beläuft es sich auf hunderte, wenn nicht tausende Nutzerinnen und Nutzer. Seit es die IEB gibt, wurden nämlich über 2.000 projektspezifische Datenauszüge für Forschende erstellt. In noch größerem Umfang genutzt werden die Datenprodukte des Forschungsdatenzentrums, die auf Standardstichproben der IEB basieren. Diese stehen dann weltweit auch externen Forscherinnen und Forschern zur Verfügung.

> Die IEB ist unglaublich flexibel.

Frau Stephan, aus Ihrer Sicht als Wissenschaftlerin: Was macht die IEB für die Forschenden so attraktiv?

Gesine Stephan: Die IEB ist unglaublich flexibel. Sie enthält Informationen zu den Erwerbsbiografien aller Menschen, die sozialversicherungspflichtig beschäftigt waren und zu allen Zeiten, in denen Menschen im Kontakt mit Arbeitsagenturen oder Jobcentern waren. Wenn wir eine Befragung durchführen und die Befragten zustimmen, können wir ihre Antworten mit den Daten verknüpfen, die in der IEB über sie vorliegen. Die IEB lässt sich auch mit anderen Prozessdaten verknüpfen. Dadurch eröffnen sich uns enorm erweiterte Auswertungsmöglichkeiten. Für uns Forschende ist die IEB eigentlich wie ein Grundnahrungsmittel. Ich vergleiche sie jetzt mal mit Kartoffeln: Kartoffeln lassen sich auch auf unendlich viele Arten zubereiten – zu einfachen Pellkartoffeln, Bratkartoffeln, Pommes Frites oder einem raffinierten Gratin. Und sie lassen sich mit allem möglichen anderen Lebensmitteln kombinieren.

Die erste intensive Nutzung des Datensatzes, das war etwa 2004, erfolgte im Rahmen der Hartz-Evaluationen (Hartz I-III), an denen Sie auch beteiligt waren. Was haben Sie damals mithilfe der IEB herausgefunden?

Stephan: Wir haben damals mit der IEB die Wirkungen von Eingliederungszuschüssen auf die Arbeitsmarktchancen untersucht. Eingliederungszuschüsse sind Zuschüsse zum Lohn, die Unternehmen für einen begrenzten Zeitraum erhalten können, wenn sie Menschen mit Vermittlungshemmnissen einstellen. Wir haben geförderte Personen mit „statistischen Zwillingen“ verglichen, also sehr ähnlichen ungeförderten Arbeitslosen, aber auch mit sehr ähnlichen Personen, die ohne Förderung eine Beschäftigung aufgenommen haben. Die Ergebnisse sprachen dafür, dass Eingliederungszuschüsse dazu beitragen können, Eintrittsbarrieren in Arbeit abzubauen und die dauerhafte Eingliederung von Personen mit Vermittlungshemmnissen zu unterstützen.

Welche wissenschaftliche Fragestellung bearbeiten Sie und Ihr Team aktuell mit der IEB?

Stephan: Ich will mal drei Beispiele nennen. Zurzeit untersuchen zwei Kolleginnen und ich, welche Menschen die Berufsberatung im Erwerbsleben (BBiE) der Bundesagentur für Arbeit in Anspruch nehmen, und ob sie sich von typischen Beschäftigten und Arbeitslosen unterscheiden. Hierfür hat uns eine Kollegin aus dem Daten- und IT-Management extra neue Forschungsdaten dazu erschlossen, welche Personen die Beratung in Anspruch nehmen. Diese haben wir dann mit der IEB verknüpft und aus der IEB auch Vergleichsgruppen von Personen gebildet, die die Beratung nicht in Anspruch genommen haben.

In einem zweiten Projekt haben ein Kollege und ich gerade analysiert, wie sich die Beschäftigung und die Beschäftigungsstrukturen im Gesundheits- und Pflegesektor in der Pandemie entwickelt haben. Und in einem dritten Projekt haben wir zu dritt gemeinsam mit externen Kooperationspartnern die IEB mit Befragungsdaten von Arbeitslosen und mit speziellen Daten zum Vermittlungsprozess verknüpft. Auf der Basis überprüfen wir, wie sich die Dauer der Arbeitslosigkeit am besten vorhersagen lässt – mithilfe von Methoden maschinellen Lernens auf Basis der IEB, durch Selbsteinschätzungen von Arbeitslosen oder aber durch die Profiling-Ergebnisse von Vermittlungsfachkräften.

> Die verstärkt mikrodatenbasierte Forschung hat eine größere Wertschätzung der Datengenerierung mit sich gebracht.

„Data ist the new oil“, hat der Economist vor einiger Zeit geschrieben. Herr Athmani, aus Sicht des IAB-Datenmanagements: Gilt das insbesondere auch für die Wissenschaft?

Athmani: Für die empirisch arbeitende Wissenschaft sind Daten schon immer ein existenzieller Bestandteil erfolgreicher Forschungsvorhaben. Gerade die methodischen Weiterentwicklungen in der Ökonometrie und der empirischen Sozialforschung in den 1990er Jahren sowie die verbesserten Computer-Kapazitäten haben zunehmend genauere und umfassendere Mikrodaten ermöglicht – aber auch erfordert. In der Konsequenz hat das zur Entwicklung der IEB geführt. Die verstärkt mikrodatenbasierte Forschung hat auch eine größere Wertschätzung der Datengenerierung mit sich gebracht. Und diese Wertschätzung hat in den letzten Jahren noch zugenommen. Die aktuellen Wirtschafts-Nobelpreise sind nicht zuletzt für herausragende Datensammlungen vergeben worden; auch die IEB hat hier eine Rolle gespielt. Wichtig ist mir aber zu betonen: Daten werden erst durch ihre gute Qualität wertvoll – und genau daran arbeiten wir im Daten- und IT-Management unermüdlich.



Bild: Kiattisak/stock.adobe.com

doi: 10.48720/IAB.FOO.20231222.01

Zitationshinweis

Christiane Keitel (2023): 20 Jahre Integrierte Erwerbsbiografien: „Für uns Forschende ist die IEB wie ein Grundnahrungsmittel“ , In: Online Magazin IAB-Forum 22. Dezember 2023, https://iab-forum.de/20-jahre-integrierte-erwerbsbiografien/, Abrufdatum: 21. May 2026

Lizenzhinweis

Diese Publikation ist unter folgender Creative-Commons-Lizenz veröffentlicht: Namensnennung – Weitergabe unter gleichen Bedingungen 4.0 International (CC BY-SA 4.0): https://creativecommons.org/licenses/by-sa/4.0/deed.de