

Ungleichheit, Armutsrisiken und Wohlfahrtsstaat. Potenziale der Datenverknüpfung von Administrativ- und Befragungsdaten für die Ungleichheitsforschung

Data Management Plan (DMP)

1 Datenerhebung und -dokumentation

1.1 Welche Daten werden Sie erheben, beobachten, generieren oder wiederverwenden?

Die verwendeten Daten sollen die Wohnbevölkerung ausgewählter und kooperierender Kantone umfassen. Auf Personenebene werden mehrere Datenquellen anhand der AHV-Nummer verknüpft. Der inhaltliche Kern der Daten wird aus den Steuerdaten der jeweiligen Kantone bestehen. Diese Daten werden dann ergänzt um Daten der Sozialhilfe (Bundesamt für Sozialversicherung BSV), Prämienverbilligungen (jeweilige Ämter der Kantone) und Ergänzungsleistungen (BSV). Zusätzlich werden über die Daten der natürlichen Bevölkerungsbewegung (BEVNAT) die Kinder und Eltern der Referenzpopulation verknüpft. Für alle so zusammengetragenen Personen werden weiter Daten der AHV- IK (Zentrale Ausgleichsstelle ZAS) sowie weitere demographische Merkmale aus der Strukturhebung, BEVNAT und der Statistik der Bevölkerung und Haushalte (STATPOP) hinzugezogen. Dieser Teil beinhaltet Variablen zur Erwerbsbiographie (aus AHV- IK), sowie Informationen zur Haushaltsstruktur, Bildung, Alter, Nationalität, familiäre Ereignisse (Ehe, Scheidung, Tod). Damit werden die Daten eine neuartige und wertvolle Grundlage zur Analyse von Armut und Ungleichheit bilden. Die Daten werden zum einen über eine hohe Fallzahl verfügen (wichtig für detaillierte Analysen) und zum anderen werden die Schwächen von Steuerdaten (z.B. das Fehlen von Sozialleistungen oder die mangelnde Abbildung von Haushaltsstrukturen) behoben.

1.2 Wie werden die Daten erhoben, beobachtet oder generiert?

Bestehende Administrativ- und Befragungsdaten werden verknüpft. Die Daten werden basierend hinsichtlich gängiger Konzepte der OECD auf ihre Tauglichkeit hin überprüft.

1.3 Welche Dokumentationen und Metadaten sehen Sie im Zusammenhang mit den Daten vor?

Um die Daten auffindbar zu machen, sollen diese dokumentiert und bei jeder Verwendung zitiert werden. Die Dokumentation wird auf der Website des Forscherteams (inequalities.ch) veröffentlicht. Damit wird es möglich sein, die Variablenbezeichnungen sowie die umfasste Population einzusehen. Zudem werden Mitarbeiter des Projektteams als Kontakt angegeben für alle Fragen die den Zugang zu den Daten betreffen. Weiterhin wird ein Projekteintrag bei FORS erstellt, der Informationen zu den Daten und deren Zugänglichkeit enthält.

2 Ethische, rechtliche und Sicherheitsfragen

2.1 Wie gestalten sich der Umgang mit und die Behandlung von ethischen Fragen?

Die Datenbasis wird sensible und umfassende Informationen zu einem grossen Teil der Schweizer Bevölkerung beinhalten. Um den Datenschutz zu wahren, ist der Zugang zu den Projektdaten daher sehr restriktiv zu gestalten. Ein Zugang zu den Daten kann nur in Absprache mit allen involvierten Stellen erfolgen und erfordert in der Regel Datenverträge mit den jeweiligen Parteien. Das umfasst die kantonalen Steuerämter, das Bundesamt für Statistik, die kantonalen Ämter für Prämienverbilligungen, das Bundesamt für Sozialversicherung und die Zentrale Ausgleichsstelle. Zur Vereinfachung streben wir eine Kooperation mit dem Bundesamt für Statistik an, bei dem die Daten in verknüpfter Version inklusive die entsprechenden Rohdaten und Skripte zur Verknüpfung und Aufbereitung lagern werden.

2.2 Wie werden der Zugriff auf Daten und die Datensicherheit verwaltet?

Müsste ausgearbeitet werden.

2.3 Wie lösen Sie Urheberrechtsfragen und Fragen des geistigen Eigentums?

Kein Thema

3 Datenspeicherung und Datenerhalt

3.1 Wie werden Ihre Daten während der Forschungsarbeiten gespeichert und gesichert?

Die Daten werden auf gesicherten Festplatten abgelegt. Dabei stützen wir uns auf die Erfahrungen der IT-Experten der Fachhochschule und der Universität.

3.2 Wie gestaltet sich Ihre Planung für den Datenerhalt?

Wie beschrieben ist vorgesehen, die Daten beim Bundesamt für Statistik (BFS) zu lagern. Auch die eigentliche Verknüpfung der Daten geschieht vor Ort beim BFS durch Mitarbeiter des Projekts (XXXXX). Da das BFS ein Eigeninteresse an den Daten hat, sollen diese dort erhalten, gepflegt und ggf. erweitert werden. Das umfasst die ausgewählten Rohdaten der einzelnen Datenlieferanten als auch die Skripte zur Aufbereitung und Verknüpfung der Daten. Als minimale Dauer der Aufbewahrung sind zehn Jahre vorgesehen.

4 Datenaustausch und Weiterverwendung von Daten

4.1 Wie und wo werden die Daten zugänglich gemacht?

Die wesentlichen Beschränkungen für die Nutzung der Daten sind rechtlicher Natur. Forschenden, die die Daten nutzen möchten, empfehlen wir eine Kontaktaufnahme mit dem Projektteam und dem BFS, damit die Forschenden beim Abklären von rechtlichen Fragen unterstützt werden können. Sofern Datenverträge mit allen involvierten Stellen geschlossen werden können, ist die eigentliche Datenlieferung unkompliziert, da der verknüpfte und aufbereitete Datensatz beim BFS fertig aufbereitet liegen wird. Zur Vereinfachung des Vorgehens ist das Projektteam bestrebt, gemeinsam mit dem BFS und den übrigen involvierten Stellen einen Standardnutzungsvertrag auszuarbeiten, der als Vorlage für die Beantragung des Datenzugangs verwendet werden kann. Die Daten werden in üblichen Formaten (namentlich R- und Stata-Datenfiles) vorliegen.

4.2 Sind bestimmte Einschränkungen erforderlich, um sensible Daten zu schützen?

siehe Punkt 4.1

4.3 Ich werde ausschliesslich digitale Archive wählen, die den FAIR Data Principles entsprechen

Ja

4.4 Ich werde digitale Archive wählen, die von einer gemeinnützigen Organisation verwaltet werden.

Ja