

# Wie erkennen wir neue Muster in der Sozialen Sicherheit?

In einem Forschungsprojekt des Fachbereichs Soziale Arbeit und des Departements Technik und Informatik der Berner Fachhochschule nutzen Forscher der BFH-TI, basierend auf Daten der Administration, komplexe statistische Analysemethoden, um das System der sozialen Sicherheit besser zu verstehen und eine Langzeitbeobachtung der Risikostruktur zu ermöglichen.



**Dr. Bernhard Anrig**  
Professor für Informatik  
BFH-TI

**Christoph Schaller**  
MSc in Engineering  
Wissenschaftlicher Mitarbeiter ICTM



**Tobias Fritschi**  
Professor für Ökonomie,  
Fachbereich Soziale Arbeit

**Oliver Hümbelin**  
Soziologe, lic. rer. soc.  
Fachbereich Soziale Arbeit

Die Schweiz besitzt ein umfassendes Netz sozialer Sicherheit. In der Erbringung von Leistungen entstehen administrative Datensätze der Arbeitslosenversicherung (ALV), der Sozialhilfe (SH) sowie der Invalidenversicherung (IV), die sich über die Sozialversicherungsnummer zeitlich verknüpfen lassen und weite Teile des Systems der sozialen Sicherheit abbilden. Data Mining umfasst statistische Methoden der Strukturerkennung und der Modellbildung. Diese haben wir systematisch auf Datenbestände der Sozialen Sicherheit angewandt.

### Drei Fragen im Fokus

Im Zentrum des Projektes mit dem Titel «Data Mining mit Administrativdaten der Sozialen Sicherheit» (2012–2014) stehen drei Kernfragen:

- Können wir mit diesen komplexen Methoden interpretierbare Verlaufsmuster identifizieren?
- Welche individuellen und kontextuellen Einflussgrößen erklären die Zugehörigkeit zu typischen Verlaufstypen?
- Lassen sich die Methoden des Data Mining in einem Analyseinstrument für die Langzeitbeobachtung nutzen?

Als Grundlage der Analysen dienen uns alle Verläufe von Personen, die in den Jahren 2005 bis 2010 in der Schweiz Leistungen der Sozialen Sicherheit (ALV, IV und Sozialhilfe) beanspruchten. Diese Daten werden vom Bundesamt für Sozialversicherungen BSV in Form des

Datensatzes SHIVALV (Sozialhilfe, Invalidenversicherung und Arbeitslosenversicherung) zur Verfügung gestellt. Die Daten wurden zudem mit Informationen zur Erwerbstätigkeit aus den individuellen Konten der AHV verknüpft, womit wir auch Erwerbsverläufe während und nach dem Bezug von Sozialleistungen rekonstruieren können. Ebenso fand eine Ergänzung der Informationen zur soziodemografischen Struktur der ALV-Beziehenden (Arbeitslosenversicherung) aus Registerdaten der Arbeitslosenversicherung statt (Arbeitsvermittlung und Arbeitsmarktstatistik / Auszahlungssystem der Arbeitslosenstellen AVAM / ASAL).

Dank der engen Zusammenarbeit des Fachbereichs Soziale Arbeit mit den Instituten RISIS (Research Institute for Security in the Information Society) und ICTM (Institute for ICT-Based Management) der BFH-TI konnten wir in der Umsetzung Kompetenzen aus Ökonomie, Soziologie, Statistik und Informatik in das Projekt einbringen und so kombinieren, dass wir aussagekräftige Resultate zuhanden des Praxispartners SECO erzielen.

### Die vier Säulen der Leistungs- und Erwerbsverläufe

Indem wir Cluster- und Sequenzanalyse verbinden, können wir die Vielzahl der beobachteten Verläufe mit 22 Verlaufsindikatoren als Grundlage für eine Verlaufstypologie zu vier Clustern gruppieren (Abb. 1). Betrachten wir das gesamte System, lassen sich folgende Gruppen unterscheiden: Für 39% der Verläufe des grössten

Clusters 1 ist eine kurze Unterstützungsphase (ALV) und ein schneller Eintritt in den Arbeitsmarkt charakteristisch. Der mit 38% fast gleich grosse Cluster 2 fasst Verläufe am Rande des Arbeitsmarktes zusammen. Diese beinhalten längere und/oder mehrere Perioden ALV sowie Phasen ohne Leistungsbezug und ohne Erwerbstätigkeit. Weitere Cluster entstanden um die Sozialhilfe (12%) und die Invalidenrente (10%).

### Wie überwacht man die Risikostruktur von Leistungsbeziehenden

In weiteren Auswertungen analysierten wir das «Risiko» eines Ausschlusses aus dem Erwerbsleben. Dabei haben wir einen Indikator gebildet, der über vier Jahre misst, wie «nahe» Erwerbsbiografien nach dem Kontakt mit dem System der sozialen Sicherheit am Arbeitsmarkt verlaufen. Basierend auf demografischen Merkmalen der Leistungsbeziehenden (Geschlecht, Zivilstand, Alter, Bildung und Wohnort) entwickelten wir ein Risikomodell zur Voraussage der Risikostruktur.

Mit diesem Prognosemodell stellen wir Merkmale von Eintrittskohorten – neu Leistungen beziehenden Personengruppen – in Bezug zu den erwarteten Möglichkeiten der Teilhabe an der Erwerbstätigkeit. So können wir die Veränderung der gesamtsystemischen Belastung der Sozialen Sicherheit gruppenweise beurteilen. In den bisherigen Arbeiten benutzten wir das Prognosemodell ausgehend von Eintritten über die ALV.

Aus den in Abb. 2 abgebildeten Werten können wir jahrweise ablesen, wie sich die Risikostruktur innerhalb von vier Jahren verändert. Höhere Werte entsprechen einer weiteren prognostizierten Distanz zur Erwerbstätigkeit. Aufgrund der über alle Neueintritte summierten erwarteten Distanz – siehe gestrichelte Linie – können wir die Veränderung der zukünftigen Gesamtbelastung des Systems nach Eintrittskohorten der sozialen Sicherheit ablesen. Aus der durchgezogenen Linie, welche die durchschnittliche Distanz zur Erwerbstätigkeit einer neu eintretenden Person darstellt, ist die veränderte Risikostruktur der Neubeziehenden ersichtlich.

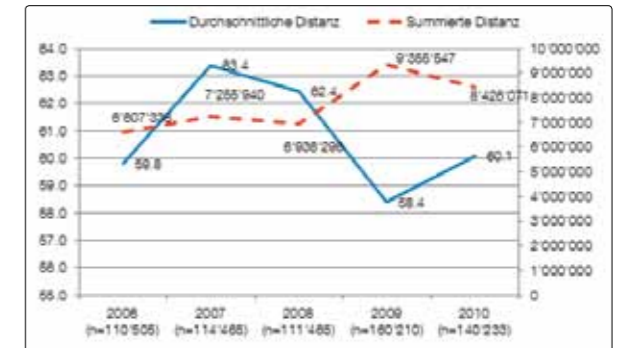


Abb. 2: Veränderung der Risikostruktur (2006 bis 2010).  
Quelle: SHIVALV 2006–2010, AVAM/ASAL 2006–2010, AHV-IK 2004–2010; Berechnungen BFH

### Im Fokus: Auswertungstool mit Excel-Abfrage

Als Teilauftrag konsolidierten wir für den Praxispartner SECO die Auswertungen ausgehend von ALV-Eintrittskohorten in einem Monitoring-Tool. Dieses basiert auf einem SQL Server Analysis Services Cube, den wir über verschiedene Software mit entsprechender Unterstützung auswerten. Über Microsoft Office Excel sind beispielsweise Abfragen mit Pivot-Tabellen möglich, dank welchen wir Indikatoren bezüglich Kalenderzeit und weiteren Merkmalen wie Region oder Demografie setzen können. Der SQL Server Reporting Service erlaubt erweiterte Darstellungsmöglichkeiten. Damit können wir Auswertungen als Diagramme und oder einfache Karten darstellen.

### Kontakt

- [bernhard.anrig@bfh.ch](mailto:bernhard.anrig@bfh.ch)
- [tobias.fritschi@bfh.ch](mailto:tobias.fritschi@bfh.ch)
- Infos: [isis.bfh.ch](http://isis.bfh.ch)

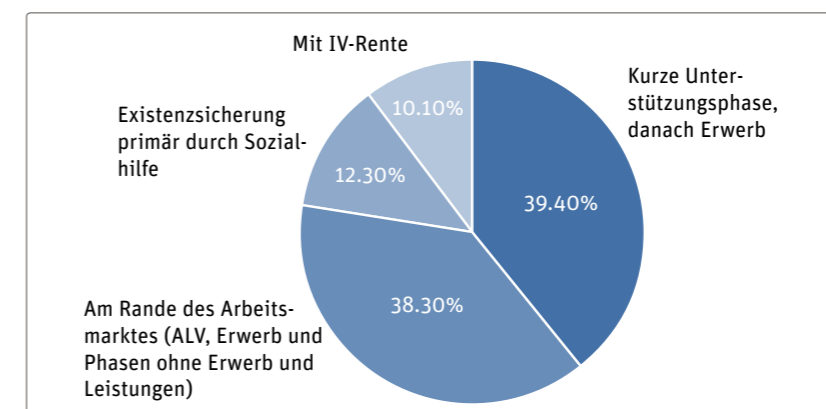


Abb. 1: Die Cluster. Quelle: SHIVALV 2006–2010, AHV-IK 2004–2010, N = 110'508; Berechnungen BFH