
Der Bedarf nach einer Treuhänderlösung für die Verknüpfung von Mikrodaten in der Bundesrepublik

Rainer Schnell und Tobias Bachteler
Universität Konstanz

1. März 2006

1 Forderungen nach einer besseren Nutzung von Mikrodaten

Auf die Notwendigkeit einer verbesserten Nutzung von Mikrodaten durch die Sozialwissenschaften in der Bundesrepublik Deutschland wird seit langem wiederholt hingewiesen. Dabei werden immer wieder zwei Aspekte betont: Zum einen wird gefordert, dass der Zugang zu den verschiedenen statistischen Datenbasen für die Wissenschaft mittels geeigneter Infrastrukturmaßnahmen erleichtert werden soll. Darüber hinaus wird empfohlen, die Möglichkeiten der exakten Zusammenführung von Mikrodaten stärker als bisher zu nutzen. So empfiehlt die „Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik“:

„Um die Kosten der Datenbeschaffung sowie die Belastung der Befragten zu senken, sollten die Anbieter der statistischen Infrastruktur ihr Augenmerk in Zukunft noch stärker als bisher auf eine volle Ausschöpfung des Informationsgehalts von vorhandenem Zahlenmaterial durch statistische Mikrodatenverknüpfung richten.“ (KVI 2001: 283)

2 Situation und Hauptproblem von Mikrodatenverknüpfungen

Im Bereich des Zugangs zu statistischen Datenbasen wurden durch die Einrichtung der Forschungsdatenzentren beträchtliche Fortschritte erzielt, im Bereich der Mikrodatenverknüpfung hat sich die Situation dagegen kaum verbessert.

Hauptgrund hierfür ist, dass derzeit keine Konzeptionen oder Vorschläge existieren, wie Mikrodaten der Sozialwissenschaften in der Bundesrepublik auf datenschutzrechtlich unbedenkliche Weise verknüpft werden könnten. Im Unterschied zu einer Datenfusion, bei der Daten von einander statistisch ähnlichen Untersuchungseinheiten verknüpft werden, werden bei einer exakten Mikrodatenverknüpfung („record linkage“) Merkmale derselben Untersuchungseinheit aus zwei verschiedenen Datenbanken oder Registern über gemeinsame Schlüssel zusammengeführt. Ohne eindeutigen Verknüpfungsschlüssel, etwa eine Identifizierungsnummer, muss zur Datenverknüpfung auf Merkmale wie Name, Geburtsdatum und Adresse zurückgegriffen werden. Damit aber setzen die Datenschutzbestimmungen in der BRD der exakten Verknüpfung von Mikrodaten aus verschiedenen sozialwissenschaftlich relevanten Datenbanken enge Grenzen, weil nur unter Wahrung der Anonymität der statistischen Einheiten die Verknüpfung datenschutzrechtlich unbedenklich wäre.

Bislang entscheiden die Aufsichtsbehörden über die Zulässigkeit einer Zusammenführung immer fallweise. Dies führt zu langwierigen und aufwändigen Genehmigungsverfahren. Entsprechend wird Record-Linkage daher seltener durchgeführt, als es angemessen wäre.

Um die wissenschaftlichen Möglichkeiten des Record-Linkage auszuschöpfen, erscheint die Entwicklung einer praktikablen Datentreuhänderlösung, die einmalig durch die Datenschutzbehörden zertifiziert werden soll, als unverzichtbarer erster Schritt.

3 Ein Treuhändermodell für Mikrodatenverknüpfungen

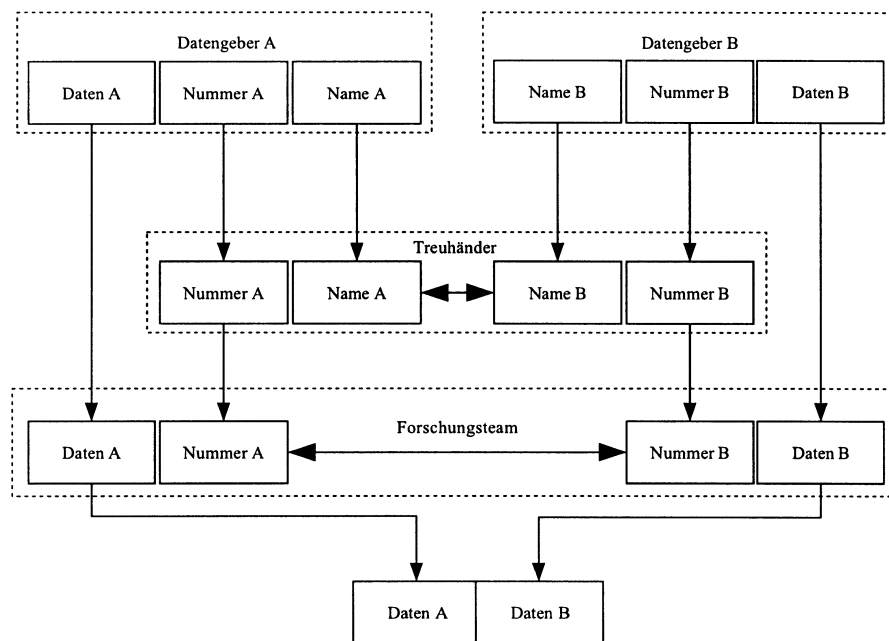


Abbildung 1: Ein einfaches Datentreuhändermodell (Quelle: Rainer Schnell, Paul B. Hill und Elke Esser (2005), Methoden der empirischen Sozialforschung, 7. Auflage, München, S.256)

Bei einem einfachen Treuhändermodell (vgl. Abbildung 1) übermitteln die beiden Datengeber dem Datentreuhänder lediglich die zur Verknüpfung benötigten Merkmale (Name, Geburtsdatum, Adresse) und eine beliebige, aber für den jeweiligen Datensatz ein-eindeutige laufende Nummer. Der Datentreuhänder verknüpft die Datenzeilen anhand der Merkmale und erhält so Paare laufender Nummern. Danach löscht der Treuhänder die Verknüpfungsmerkmale und übermittelt der Forschungsgruppe lediglich die Paare laufender Nummern. Die Forschungsgruppe kann die von den Datengebern zuvor übermittelten

Sachdaten zusammenführen, ohne dass die identifizierenden Merkmale bekannt sind.

Nach juristischer Meinung sollte eine solche Treuhänderstelle mit einem Zeugnisverweigerungsrecht ausgestattet sowie die dort gespeicherten Daten mit einem Beschlagnahmeverbot gegen staatlichen Zugriff geschützt sein. Dafür kämen dann nur Notare oder eventuell die Justitiare von Universitäten in Frage.

4 Durchführung eines Workshops zu einem Datentreuhändermodell

Es wird daher angeregt, einen Workshop aller Interessierten zu Datentreuhändermodellen durchzuführen. Hierbei sollen durchgeführte und geplante Beispiele vorgestellt und diskutiert werden. Der Workshop soll die zu erwartenden Probleme klären und dazu beitragen, die weiteren Schritte zur Implementierung eines Modells festzulegen.

Eingeladen werden sollten Vertreter der

- Forschungsdatenzentren
- Krebsregister
- Disease-Management-Programm-Evaluationen
- in diesem Bereich aktiven Sozialforschungsinstitute
- in diesem Bereich aktiven inhaltlichen Forschungsgruppen der Sozialwissenschaften (z.B. Migrationsforschung und Kriminologie)
- aktiven Record-Linkage-Forschungsgruppen

Die Organisation des Workshops könnte der RatWSD zusammen mit dem Zentrum für Quantitative Methoden und Surveyforschung an der Universität Konstanz übernehmen.