

Workshop: Scientific Use Files aus verknüpften Sozialdaten

Veranstalter: SOEP, Ort: DIW Berlin, Zeit: 17. und 18. November 2015

Thematische Schwerpunkte:

- Dienstag: Datenzugang, Datenschutz und Scientific Use Files (SUF) - derzeitige Praxis und Zukunftsaussichten
- Mittwoch: Angriffsszenarien und Deanonymisierungsmöglichkeiten auf dem Prüfstand

Dienstag, 17.11.2015 ab 12 Uhr

12:00 - 13:00 Registrierung, Imbiss

1 Einführung in die Thematik

13:00 - 13:30

Jan Goebel, Leiter des Forschungsdatenzentrums des SOEP

2 Vorträge: Datenschutz und Forschungsdatenzugang bei Sozialdaten

13:30 - 15:30: jeweils 30 min für die Vorträge und die Diskussion

2.1 Spannungsfeld Sozialdatenschutz und Nutzung von Sozialdaten durch die Wissenschaft - Handhabung bei der DRV

Herr Gerold, Mitarbeiter im Datenschutzreferat der DRV Bund

2.2 Spannungsfeld Sozialdatenschutz und Nutzung von Sozialdaten durch die Wissenschaft - Handhabung bei der BA

Elisabeth Roß, Mitarbeiterin des IAB im Wissenschaftsmanagement, zugleich Datenschutzkoordinatorin

Manfred Antoni, Mitarbeiter des Forschungsdatenzentrums (FDZ) der Bundesagentur für Arbeit im IAB

2.3 Spannungsfeld Sozialdatenschutz und Nutzung von Sozialdaten durch die Wissenschaft aus Sicht der Aufsichtsbehörde BMAS

Simone Solka, Mitarbeiterin im Referat Ila2 des Bundesministeriums für Arbeit und Soziales

2.4 Diskussion

15:30-16:00 Kaffeepause

3 Anonymisierungsmaßnahmen zur Erzeugung von Standard-SUFs und in verschiedenen Verknüpfungsprojekten

16:00- 17:30

3.1 Synopse der Anonymisierungsmaßnahmen in verschiedenen Standard-SUFs und der Empfehlungen von Müller et al. (1991)

Klaudia Erhardt, Mitarbeiterin SOEP

~ 30 min

Mikrozensus und Mikrozensus Regionalfile - SIAB regional (Bundesagentur f. Arbeit) - Rentenversicherungsdaten der DRV - Erwachsenenkohorte (Startkohorte 6) des NEPS - DZA Alterssurvey und DZA Freiwilligensurvey - Sozio-oekonomisches Panel

3.2 Darstellung der SUF-Situation, Probleme und Lösungen für verknüpfte Datensätze

~ 60 min

Vorstellung der SUF-Situation bzw. des aktuellen Stands für den verknüpften Datensatz entlang von thematischen Gesichtspunkten

- Share-RV (Christin Czaplicki)
- BASiD (Tanja Mika)
- NEPS-SC6-ADIAB (Manfred Antoni)
- LidA (SUF-Bereitstellung nicht geplant) (Stefanie March)
- SOEP-IAB Mig-Stichprobe (Klaudia Erhardt / Parvati Trübswetter)
- Evaluations-Datensatz des IZA (Nikos Askitas)

4 Diskussion und Zusammenfassung der heutigen Ergebnisse

17:30 - 18:00

Am Abend ab 19:30 optionales gemeinsames Abendessen
(auf eigene Kosten der Teilnehmenden)
im Restaurant "Mikoto" am Schlesischen Tor

Mittwoch 18.11.2015 ab 9:30

5 Vortrag: Weitergabe von Forschungsdaten zwischen Risiko und Ungewissheit

Rainer Lenz, Prof. für Angewandte Statistik und Mathematik an der HTW Saarland

9:30 - 10:30 (inkl. Diskussion)

10:30-11:00 Kaffeepause

währenddessen Umstellen der Tische zu Arbeitsecken für kleinere Gruppen

6 Diskussion mehrerer Themen in kleineren Gruppen auf Basis von im Vorfeld verschicktem Input und der bisherigen Erkenntnisse des Workshops

11:00 - 14:00, dazwischen 1/2 Stunde Pause und Mittagessen (letzteres teilw. während der Diskussion).

Die vorgeschlagenen Diskussionsthemen sind:

- Big Data als Anonymisierungsrisiko für verknüpfte Sozialdaten
- Realistische Angriffsszenarien zur Deanonymisierung von verknüpften Survey- und Sozialdaten
- Vertrauen ist gut, Kontrolle ist besser? wie zuverlässig ist die Wissenschaft als Datenschutzpartner? Was kann getan werden, um das Vertrauen der FDZs in die externen ForscherInnen zu stärken?

Zugunsten der Diskussion Aufteilung in 3 Kleingruppen.

Die Zeit reicht nicht aus, dass jede Gruppe alle 3 Themen diskutiert. Falls nicht eines der Themen mangels Interesse wegfällt, diskutiert jede Gruppe 2 der 3 Themen. Da es 3 Kombinationen gibt, kann jede Gruppe eine andere Themenkombination wählen.

Erforderlich: "Gruppenpaten", welche die Ergebnisse festhalten und später vorstellen.

6.1 Vorstellung des Konzepts, Organisatorisches und kurze Einführung in die drei Themen

11:00 -11:30

Klaudia Erhardt

6.2 Die Gruppen diskutieren ihr erstes gewähltes Thema

11:30 - 12:30 Diskussion

12:30-13:00 Mittagspause

Eine halbe Stunde ist ziemlich kurz für alles, was normalerweise in den Pausen erledigt wird. Da einige der extern angereisten TeilnehmerInnen zeitig zum Zug müssen, war es nicht möglich, eine ganze Stunde für die Mittagspause vorzusehen.

Daher ist geplant, den Mittagsimbiss bei Bedarf in die folgende Diskussion auszudehnen - ausnahmsweise sei erlaubt, mit vollem Mund zu reden ☺.

6.3 Die Gruppen diskutieren ihr zweites gewähltes Thema

13:00 - 14:00 Diskussion

14:00 - 14:15 Wieder zum Plenum versammeln / zurückstellen der Gruppentische

7 Vorstellen der Ergebnisse der Diskussionen und sie in Zusammenhang bringen zur Möglichkeit einer faktischen Anonymisierung von verknüpften Daten

14:15 - 15:15

8 Bilanz des Workshops, Ausblick: wie geht es weiter mit den SUFs für verknüpfte Daten?

15:15 - 15:30

plus 30 min Puffer, da der heutige Zeitplan sehr straff ist

16:00 Ende des Workshops
