

Konferenz für Sozial- und Wirtschaftsdaten  
20./21. Februar 2014, Berlin

Gute Forschung braucht gute Daten – aber bitte anonymisiert!

# Big Data, Amtliche Statistik und der Datenschutz

**Peter Schaar**

Europäische Akademie für  
Informationsfreiheit und Datenschutz  
[www.eaid-berlin.de](http://www.eaid-berlin.de)

# Das klassische Modell

- Aufgabe
- Lösungsweg (Algorithmus)
- Daten ?

# Das Big Data Modell

- Daten
- Korrelation
- Erkenntnisse ?
- Prognose !

# Statistik als Big Data?

## Haushaltsbefragung auf Stichprobenbasis zum Zensus 2011

Stichtag: 9. Mai 2011

Platzhalter für  
Etikett-/Fragebogen-Nr.

Platzhalter für  
Barcode/Fragebogen-Nr.  
2701000001076

### Zweck der Erhebung

Die Haushaltsbefragung dient einerseits der Qualitätssicherung der registergestützt ermittelten Einwohnerzahl. Andererseits dient die Haushaltsbefragung auch der Erhebung von Zensusmerkmalen, die nicht aus Verwaltungsregistern gewonnen werden können.

Es besteht Auskunftspflicht, mit Ausnahme zu Frage 8. Die Beantwortung der Frage 8 ist freiwillig.

Den Fragebogen können Sie auch im Internet ausfüllen.  
Wir haben für Sie unter [www.zensus2011.de](http://www.zensus2011.de) bereits alles vorbereitet.

**online**

Ihre Fragebogennummer: 2701000001076 Ihr Aktivierungscode: zWkLvccGprwa

Für jede Person des Haushalts ist je ein Fragebogen auszufüllen.

Vielen Dank für Ihre Unterstützung.

Rechtliche Hinweise entnehmen Sie den Seiten 9 und 10 dieses Fragebogens.

### Gehen Sie wie folgt vor:

- Beantworten Sie die Fragen der Reihe nach.  
4. Zahlen tragen Sie bitte rechtsbündig ein.
- Kreuzen Sie bitte, sofern nicht anders angegeben, für jede Frage nur eine Antwort an.  
Anzahl der Personen .....
- Überspringen Sie Fragen nur dann, wenn hinter dem von Ihnen angekreuzten Kästchen der Hinweis „Weiter mit Frage ...“ steht.  
5. Text tragen Sie bitte in Druckbuchstaben ein.  
Vorname/-n:   
Nachname:
6. Falls Sie eine Antwort korrigieren müssen, nehmen Sie die Korrektur deutlich sichtbar vor.  
Ja .....  Nein .....

### Persönliche Angaben

1 Vorname/-n:

Nachname:

2 Straße, Hausnummer:

PLZ, Ort:

3 Telefonnummer:

4 Welches Geschlecht haben Sie?  
Männlich .....   
Weiblich .....

5 Wann wurden Sie geboren?  
Tag  Monat  Jahr



# Verbot von Persönlichkeitsprofilen

„Mit der Menschenwürde wäre es nicht zu vereinbaren, wenn der Staat das Recht für sich in Anspruch nehmen könnte, den Menschen zwangsweise in seiner ganzen Persönlichkeit zu registrieren und zu katalogisieren, sei es auch in der Anonymität einer statistischen Erhebung, und ihn damit wie eine Sache zu behandeln, die einer Bestandsaufnahme in jeder Beziehung zugänglich ist.“

Bundesverfassungsgericht  
Beschluss v. 16.12.1969 („Mikrozensusentscheidung“)

# Recht auf informationelle Selbstbestimmung

„Das Grundrecht gewährleistet ... die Befugnis des Einzelnen, grundsätzlich selbst über die Preisgabe und Verwendung seiner persönlichen Daten zu bestimmen. Einschränkungen dieses Rechts auf "informationelle Selbstbestimmung" sind nur im überwiegenden Allgemeininteresse zulässig.“

Bundesverfassungsgericht  
Urteil v. 15.12.1983 („Volkszählungsurteil“)



# „reguläre“ vs. statistische Datenverarbeitung

„Bei den verfassungsrechtlichen Anforderungen ... ist zu unterscheiden zwischen *personenbezogenen* Daten, die in individualisierter, nicht anonymer Form erhoben und verarbeitet werden, und solchen, die für *statistische Zwecke* bestimmt sind.“

Bundesverfassungsgericht  
Urteil v. 15.12.1983 („Volkszählungsurteil“)

# Statistische Datenverarbeitung „Schranken“

„Bei der Datenerhebung für *statistische Zwecke* kann eine enge und konkrete Zweckbindung der Daten nicht verlangt werden. Der Informationserhebung und Informationsverarbeitung müssen aber innerhalb des Informationssystems zum Ausgleich entsprechende *Schranken* gegenüberstehen.“

Bundesverfassungsgericht  
Urteil v. 15.12.1983 („Volkszählungsurteil“)

# Statistikgeheimnis

## - § 16 BStatG -

„(1) Einzelangaben über persönliche und sachliche Verhältnisse, die für eine Bundesstatistik gemacht werden, sind ... geheimzuhalten. ...“

# Übermittlung an die Wissenschaft - § 16 BStatG -

„(6) Für die Durchführung wissenschaftlicher Vorhaben dürfen... Einzelangaben an Hochschulen oder sonstige Einrichtungen ... unabhängiger wissenschaftlicher Forschung übermittelt werden, wenn die Einzelangaben nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft zugeordnet werden können und die Empfänger Amtsträger, für den öffentlichen Dienst besonders Verpflichtete oder Verpflichtete ... sind.“

# Re-Identifizierungsverbot

## - § 21 BStatG -

Eine Zusammenführung von Einzelangaben aus Bundesstatistiken oder solcher Einzelangaben mit anderen Angaben zum Zwecke der Herstellung eines Personen-, Unternehmens-, Betriebs- oder Arbeitsstättenbezugs außerhalb der Aufgabenstellung dieses Gesetzes oder der eine Bundesstatistik anordnenden Rechtsvorschrift ist untersagt.

# Goldenes Big Data- Zeitalter?

- Nicht bloß wissenschaftlicher Ansatz
- Selbstbestimmung
- Macht
- Missbrauch

# Big Data - datenschutzfreundlich?

- Beschränkung auf vorhandene Daten (keine Vorratsdatenspeicherung)
- Verwendung anonymisierter/pseudonymer Daten
- Transparenz der Verarbeitung/Bewertung
- Ausschluss sensibler Daten
- Keine automatisierten Entscheidungen

# Amtliche Statistik - Big Data

- gesetzliche Grundlage
  - große Datenmenge
  - Stichproben- und Vollerhebung
  - bestimmte Merkmale
  - möglichst exakt
  - definierte Verarbeitung
  - Trennung von operativer DV
  - Keine Rückschlüsse auf Einzelfälle
  - Verbot der Re-Identifizierung
- ohne gesetzliche Grundlage
  - riesige Datenmenge
  - möglichst „alle“ Daten
  - beliebige Merkmale
  - weniger exakt
  - Korrelation beliebiger Merkmale
  - Verbindung zu operativer DV
  - Rückschlüsse auf Einzelfälle möglich und gewollt
  - Re-Identifizierung möglich



# Diskussionpunkte

- Macht Big Data die amtliche Statistik überflüssig?
- Verwendung von Big Data-Methoden in der Statistik?
  - Register-/Sekundärerhebungen
  - Forschungsdatenzentren
  - Georeferenzierung?
  - Melderegister als statistischer Datenpool?