

WIRTSCHAFTSSTATISTISCHE PANELDATEN UND FAKTISCHE ANONYMISIERUNG

4. Konferenz für Sozial- und Wirtschaftsdaten
Wiesbaden, 20. Juni 2008
Dr. Markus Zwick

Faktische Anonymisierung Wirtschaftsstatischer Einzeldaten (FAWE)

- Projektzeitraum: 2001 bis 2005
- Erstmalige Anonymisierung wirtschaftsstatischer Einzeldaten
- Scientific Use Files für die Kostenstrukturerhebung, Einzelhandelsstatistik, Umsatzsteuerstatistik, IAB-Betriebspanel

Ronning et al (2005), Handbuch zur Anonymisierung wirtschaftsstatischer Einzeldaten, Statistik und Wissenschaft, Band 4

Wirtschaftsstatistische Paneldaten und faktische Anonymisierung (FAWE-Panel)

- Projektzeitraum: Januar 2006 bis Dezember 2008
- Erstellung von Paneldatensätzen und Verbesserung des Analysepotentials durch Erweiterung der Datensätze
- Bewertung der Auswirkung von Anonymisierungsmethoden auf Datenschutz und Analysepotential
- (Weiter)entwicklung von Methoden zur Anonymisierung von Paneldaten
- Faktische Anonymisierung der Projektstatistiken

Projektstatistiken FAWE-Panel

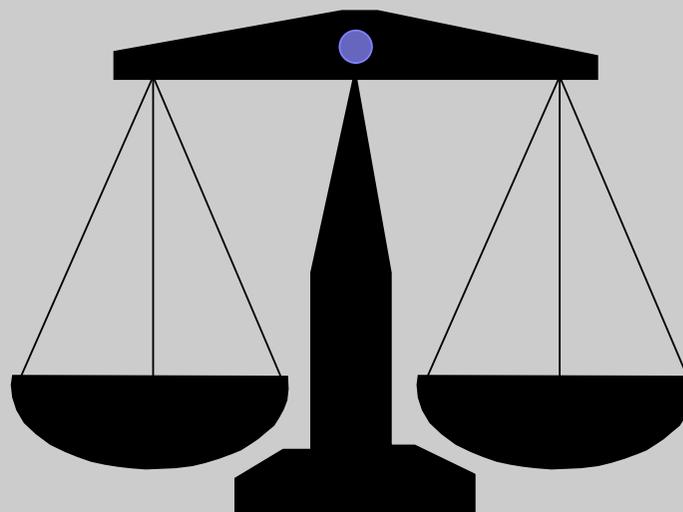
- **Kostenstrukturerhebung der Unternehmen des Verarbeitenden Gewerbes (1995-2004)**
- **Monatsberichte für Betriebe im Verarbeitenden Gewerbe (1995-2004), ergänzt um die Investitions- und Kleinbetriebserhebung im Verarbeitenden Gewerbe**
- **Umsatzsteuerstatistik (2001-2004)**
- **IAB Betriebs-Panel (1993-2004)**

Projektpartner



Zielkonflikt mit Verfassungsrang

Schutz der
Merkmalsträger

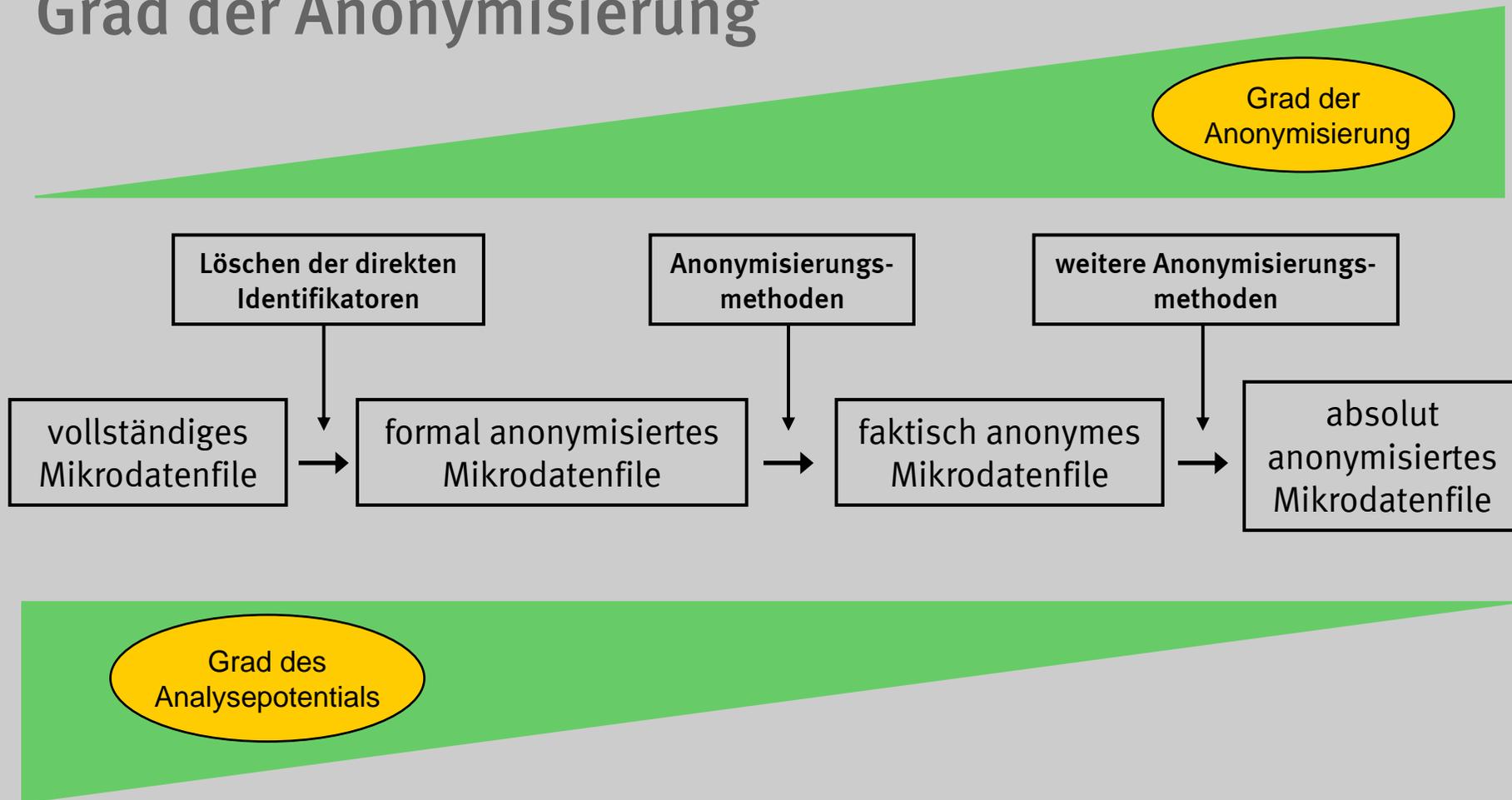


Bestmöglicher Erhalt
des Analysepotentials

Anonymisierte Einzeldaten (Mikrodaten)

- formal anonymisierte Einzeldaten
vertrauliche Einzeldaten
- absolut anonymisierte Einzeldaten (Public Use Files)
- faktisch anonymisierte Einzeldaten (Scientific Use Files)

Grad der Anonymisierung

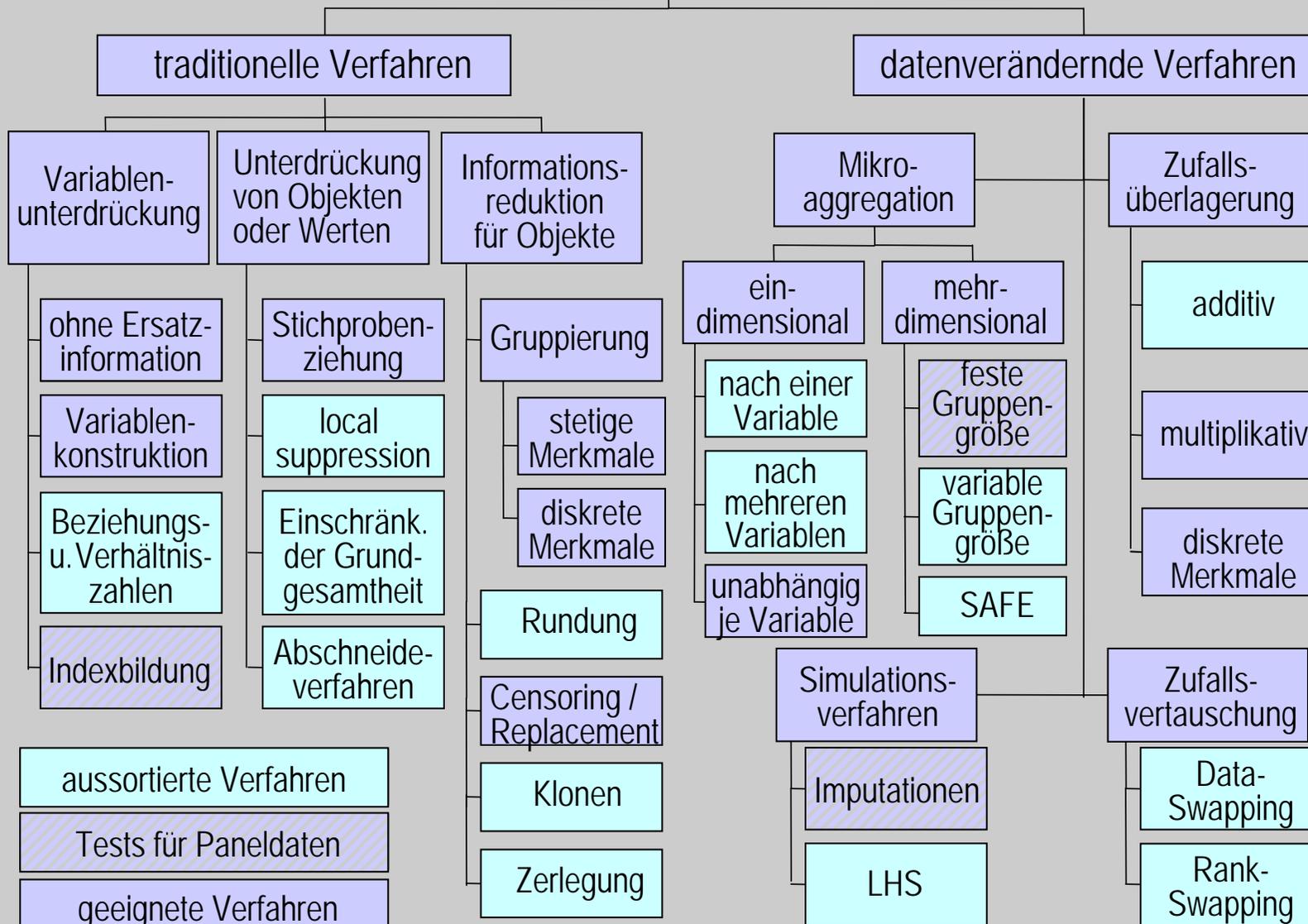


Anonymisierungsstrategien bei wirtschaftsstatistischen Daten

Mikrodaten von Unternehmen unterscheiden sich in wesentlichen Punkten von denen mit Personenbezug

- wesentlich kleinere Grundgesamtheiten
- Stichprobenauswahlsätze sind bei Unternehmenserhebungen wesentlich größer
- starke Zusammenhänge zwischen einzelnen Erhebungen
- Zusatzwissen im Bereich der Wirtschaft ist in wesentlich höherem Umfang vorhanden
- der Nutzen einer Deanonymisierung ist deutlich höher

Anonymisierungsverfahren



bisher ungelöste Probleme

- deutliche Informationsreduktion bei Scientific Use Files zur Off-Site Nutzung
- datenverfälschende Verfahren werden überwiegend abgelehnt
- Korrekturverfahren bisher eingeschränkt verwendbar
- klarer Trend zur Nutzung der On-Site Files und der kontrollierten Datenfernverarbeitung
- Angebote der FDZ bisher nur eingeschränkt für diese deutliche Verlagerung der Nachfrage ausgerichtet

Lösungsansätze

- Entwicklung anonymer Datenstrukturfiles die semantisch verwertbar sind und es ermöglichen, syntaktisch und inhaltlich fehlerfreie Auswertungsprogramme zu schreiben
- Entwicklung und Bewertung von Verfahren der standardisierten und vollständig automatisierten Ergebniskontrolle.
- Simultane Betrachtung der Anonymisierung von Mikrodaten und der Ergebniskontrolle.

Ziel: Remote Access

Eine informationelle Infrastruktur für das ‚E-Science Age‘

Auf dem Weg zum ‚Remote-Access‘

Verbesserung der Angebote der kontrollierten Datenfernverarbeitung durch Datenstrukturfiles und automatisierte Ergebniskontrolle

Vision 2027



- Mikrodaten der verschiedenen Datenproduzenten sind umfassend verknüpft
- Zugriff vom Arbeitsplatz des Wissenschaftlers auf die Mikro- wie Metadaten 24/365
- die entwickelten technischen und rechtlichen Verfahrensweisen stehen im vollen Einklang mit dem Datenschutz

Vielen Dank für Ihre Aufmerksamkeit