

■ **Tanja Schmidt**  
**Gemeinsame Nutzung von Ressourcen für  
die sozioökonomische Berichterstattung  
»soeb«**

**Workshop des WissGrid Fachberater-Teams  
Göttingen SUB, 19. Januar 2011**

**SOFI**

Soziologisches Forschungsinstitut  
an der Universität Göttingen

Tanja Schmidt

 **soeb.de**

# ■ Soziökonomische Berichterstattung »soeb«

## ■ Berichterstattung zur sozioökonomischen Entwicklung Deutschlands

- Umbruch in Wirtschafts- und Lebensweise
- Auswirkungen auf Wohlfahrt und Ungleichheit
- Wohlfahrtsmessung – BIP ist kein Wohlfahrtsmaß

## ■ Forschungsverbund sozialwissenschaftlicher Institute mit Mitteln der Forschungsförderung (BMBF)

- Neue Brücken zwischen Sozialforschung und Sozialberichterstattung
- Übertragung von Forschungsergebnissen in Konzepte und Indikatoren für Dauerbeobachtung

## ■ Mehr Information: [www.soeb.de](http://www.soeb.de)

# ■ »soeb« als Anwendungsfall für eine virtuelle Forschungsumgebung

## ■ Sozioökonomische Berichterstattung als besonderer Anwendungsfall

- Viele Verbundpartner/innen
- Nutzung vieler verschiedener Mikrodatensätze zum Teil durch mehrere Bearbeiter/innen
- Datenkonventionen für replizierbare Indikatoren
- Relativ hoher Abstimmungsbedarf bei Syntaxlösungen
- Komplexe Vernetzungs- und Zulieferbeziehungen

## ■ Anforderungen und Lösungen lassen sich verallgemeinern

- Sozialwissenschaftliche Forschungsabläufe und –schritte sind typisch
- Probleme der Datenbereitstellung und Nutzung sind repräsentativ

# ■ kollaborative Datenauswertung in der »soeb«

## ■ Arbeitsteilige Arbeit an Datensätzen

- Aufbereitung von Arbeitsdateien, Data Matching (Beispiele SOEP, LIAB)
- Gemeinsame Erstellung modularer Auswertungssyntax
- Transfer spezieller Daten- und Methodenexpertise
- Nutzung von Daten, Syntax in verschiedenen Arbeitspaketen
- Konvertierung in andere Statistikprogramme
- Archivierung von Outputs, Übertragung in Excel

## ■ Nachnutzbarkeit

- für spätere Arbeitsphasen: Aktualisierung, lange Reihen
- für berechnete Dritte

# ■ Probleme im kollaborativen Arbeitsprozess

## ■ Kein gemeinsamer Zugriff auf Ausgangsdaten

- Onsite-Files, SUF nicht identisch
- Probleme bei explorativer Datenanalyse, Fehlersuche, Übernahme von Syntaxmodulen

## ■ Gemeinsame Nutzung generierter Arbeitsdateien kann Datenschutzbelangen und Nutzungsregeln der Datenhalter widersprechen

## ■ Beschränkte PC-Speicher- und Rechenkapazität

- bei großen Datensätzen (z.B. IEBS; LIAB)
- bei komplizierten Längsschnittverfahren (z.B. Sequenzanalysen, Optimal Matching Analyse)

## ■ Austausch von Dateien und Arbeitsergebnisse problematisch

# ■ Hauptfunktionen in einer virtuellen Forschungsumgebung (1)

## ■ Datenanbieterzugriff

- Zugang zu Ausgangsdaten von FDZ (SUF)
- Fernrechnen und Ablage von Daten zur langfristigen Archivierung

## ■ Datenverwaltung

- Forschungsdatensätze, Syntaxdateien und Metadaten
- Suchfunktionen für alle Datenarten, Datenversionierung, Verknüpfungen
- Individuelle Nutzungsrechte und –Zeiträume durch Zutritt zu virtuellen Organisationen

## ■ Datenbearbeitung

- Editieren von Metadaten und Syntaxdateien

## ■ Datenvergleich

- Werkzeuge zum Vergleich von Syntaxdateiversionen

# ■ Hauptfunktionen in einer virtuellen Forschungsumgebung (2)

## ■ Datenverarbeitung

- Ausführen von statistischen Berechnungen an Forschungsdaten

## ■ Konfiguration

- System-, Installations- und Nutzer/innen Einstellungen

## ■ Verwaltung

- Verwaltung von Benutzergruppen, Rollen und globalen Rechten

## ■ Kollaboration

- Werkzeuge für die Zusammenarbeit im Verbund

## ■ Publikation

- Veröffentlichung von Forschungsergebnissen im Forschungsverbund sowie auf Webseiten

# ■ Kollaborative Datennutzung aus Sicht des Datenschutzes – Probleme und Lösungen

## ■ Unproblematisch:

- Ablage von Metadaten, gemeinsamer Zugriff auf Syntaxdateien in der Forschungsumgebung

## ■ Probleme mit Ausgangsdaten in virtuellen Organisationen:

- Personelle Nutzungsbeschränkung, Genehmigungsvorbehalt
- Inhaltliche Nutzungsbeschränkung (Zweckbindung)
- Zeitliche Nutzungsbeschränkung (Befristung, Löschfristen)

## ■ Mögliche technische Lösungen in virtueller Arbeitsumgebung

- Zugriffskontrolle durch Rechteverwaltung für virtuelle Organisationen: zertifikatsbasierte Authentifizierung (Grid Security Infrastructure: PKI)
- Transparenz der Datennutzung durch gemeinsame Dokumentation des »Lifecycles« (Protokollierung, Historie)
- Rechteverwaltung gewährleistet rollenbasierte Rechte und Löschfristen



# ■ Wie die virtuelle Forschungsumgebung Arbeitsweisen von Forscher/inne/n verändert

## ■ Arbeit im Verbund ...

- ... bleibt dezentral und individuell
- ... wird durch technische Vernetzung transparenter
- ... setzt künftig Bereitschaft zu »syntax sharing« voraus
  - Syntax an (teil-)öffentlichem Ort

## ■ Verhaltensanforderungen an Verbundpartner/innen ...

- ... werden technisch »hinterlegt« und formalisiert
- ... Nutzer/inne/n werden bei Einhaltung von Standards wirksam unterstützt

## ■ Verbundkoordination ...

- ... wird anspruchsvoller, um technischen Support erweitert

## ■ Entwicklung der virtuellen Forschungsumgebung in engem Kontakt mit Forschungspraxis

# ■ Dokumentation

## ■ Berichterstattung zur sozioökonomischen Entwicklung Deutschlands

- <http://www.soeb.de/virtaug.php>