A tale of two eagles: comparing and contrasting the social science data landscapes in the USA and Germany

Stefan Kramer (German Institute for Economic Research - DIW Berlin) & Denis Huschka (German Data Forum - RatSWD)
Washington, DC, June 7, 2012
Agenda: a short look at a few corners of the data landscapes in Germany & the USA:

- Statistical agencies and government data
- The state of data librarianship
- Attitudes towards the national census
- Restricted-access research data centers (RDCs)
- National data policies and their coordination

... and in closing: Potential areas of future collaboration – between these countries and others
Statistical agencies and government data

USA:
- Different federal gov’t. agencies have their own statistical/data branches (e.g., Bureau of Justice Statistics, U.S. Bureau of Economic Analysis), no central stat. agency
- States have such units too (e.g., NY State Data Center), but draw a lot on federal data collection activities (esp. Census)
- Data.gov aims to pull information together, but the decentralized nature of data sources complicates things

Germany:
- Federal statistical office (Statistisches Bundesamt - DeStatis) centralizes data gathering/processing activities
- The 16 state statistical offices (Statistische Landesaemter) undertake significant data gathering activities (incl. for the nationwide Census)
- „Open government/data“ movement in early stages (at regional level first, e.g.: http://daten.berlin.de/), nationwide planned for 2013 (http://www.daten-deutschland.de/startseite.html)
- Data from the statistical offices = +€€€!
USA:
- Data librarians at many of the larger research universities; in some cases, for decades; many more recent appointments
  - Early boost by U.S. Census data releases after 1990 -> gov’t. doc. librarians
  - Latest drivers incl.: empirical research requirements for undergrad. students; data mgmt. planning req’s from funders impact researchers

Germany:
- Data librarianship not yet as full-fledged positions, with very few exceptions
  - No response to two InetBib inquiries in 2009 & 2011
  - difficulties of matching up with job descriptions & rankings in public service?

Structured, sufficient training for developing the profession still sparse in both countries

See also “Data Librarianship: Past, Present, Future, Challenges, Opportunities”: http://hdl.handle.net/1813/19484
Attitudes towards the national census

USA:
- Census of Population & Housing since 1790 – constitutionally required for seat allocation in House of Rep’s.
  - Fair to say: a certain number of citizens don’t like it, but generally accepted?

Germany:
- Volkszählung („count of people“) every 5-9 years until 1970 in West Germany (FRG), once a decade in East Germany (GDR)
- FRG: 1981 first count postponed to 1983, then cancelled after public protests and court order
- Volkszählung 1987 happens against backdrop of protests, concerns about privacy, government snooping
- No more until Zensus 2011 (mixed-methods) as part of EU-wide requirement
Restricted-access governmental research data centers (RDCs)

USA:
• Census Bureau RDCs in 13 locations (more to come)
• Same data available in the different centers

Germany:
• Forschungsdatenzentren (RDCs) of the federal (3 locations) and state (15 locations) governments
• Same data available in the different centers
USA:

- Key report (2009): *Harnessing the power of digital data for science and society*
  - “…the combined effort of representatives from 22 federal agencies working together under the Interagency Working Group on Digital Data”
  - Influenced the National Science Foundation’s Data Management Plan requirements that were announced in May 2010
- Noteworthy for their brevity: [Data Policy Statements of data.gov](#)

Germany:

- Data protection concerns and laws abound – challenging for discussions on data sharing and publishing
  - RatSWD (Rat für Sozial- und Wirtschaftsdaten = German Data Forum) est. in 2004 by Federal Ministry of Education and Research (BMBF) advises federal & state gov’ts. on research data infrastructure
    - [Working Group on Future Data Access](#)
    - [Accreditation and support of Research Data Centers](#)
Potential areas of future collaboration – between these countries and others

- Educating the next generation of librarians to “do data” – across national boundaries
- Developing best practices for **persistent identification** of data (following *Metadata and Persistent Identifiers for Social and Economic Data workshop*)
- Advocating for global “Open Data” from public sources - in cost (trivial or none), discoverability and access (as easy as possible)
- Establishing methods for linking research publications, research data, and researchers (e.g. with [ORCID](https://orcid.org))
- Establishing coordination activities for structuring and operating research data centers
Thank you for your attention!

- Questions or comments?
- Stefan Kramer
  SKramer@DIW.de

Denis Huschka
DHuschka@RatSWD.de