A European Research Infrastructure for Social Science and Humanities Data

Developments in cross-European data sharing

Access to data: The state of play

- Focus of presentation on both *microdata* - censuses/surveys/registers - and *macrodata* - NGO/NSI outputs

- I will *not* be covering realtime data, nor commercial/customer tracking databases. And just a single mention of point-to-point data

- I will *not* be focusing on issues of comparability – these problems are taken as read!

- I will focus on re-use of data for commercial purposes

- Aim to inform 'Where we are now'
  – Hilary to shed some light on one future development!

Microdata sources

- **Censuses**
  - IPUMS-I

- **Surveys**
  - Eurostat
  - CESSDA portal
  - ESDS International
    - Guide to freely available data
    - Cross-national research resource

Macrodata sources

- ESDS International maintains a pretty exhaustive guide at http://www.esds.ac.uk/international/resources/guide.asp#freedata

- Largely, provision of data at country level

Microdata records with detailed geographies attached

- Will exists to maximise use of **census data**, including commercial access – see excellent overview by Denise Lievesley

- In practice, limited access to census microdata for non-scientific use!
  - But note for UK: Non-academic access to SARs and Longitudinal Study

- **Survey data** – more liberal access regimes but coarse geographical detail
  - UK example: commercial users have access to EUL (GOR, NUTS1) but not 'Special Licence' (NUTS2-5) data

Fine-grained macrodata sources?

- **NGOs**
  - Free online access to Eurostat (selection of data to NUTS2/3; point-to-point transport data)
  - Other freely available data via: e.g. undata
  - Others mainly by subscription: e.g. SourceOECD

- **NSIs**
  - UK: Neighbourhood Statistics, NOMIS (labour market); Census
  - Patchy (subject coverage, level of geographical detail) across other European NSIs but worth taking a look at each NSI, e.g. Austria
  - N.b. difficulty in putting together a Europe-wide or multi-nation dataset

Key messages

- Limits to geographical detail available – cross-national, cross-regional comparisons much easier to make than, e.g., ward-level comparisons

- Access restrictions/delays/charges for commercial use need to be considered

- Finding/accessing European data at fine geographical detail – sadly, there is no golden bullet!

What is CESSDA?

- Council of European Social Science Data Archives

- Currently, informal group of 20 member archives

- Goal: to facilitate pan-European data discovery & sharing

The CESSDA ERI: a legal entity to the promote access to European SSH data

- Central co-ordinating body for SSH data resources
- Governance Structure
- Sustainable funding
- Contractual agreements with service providers

- In 2007, 20 members jointly held over 25000 datasets

CESSDA mission

- Provide better, enhanced knowledge of and access to wider range of social science data
  - Improve researcher/learner experience
  - Enable quality research / training

- Maximise stakeholder investment

- Enhance transfer of skills, knowledge

Functions of the centre

- Development and maintenance of quality data collections (central and virtual) through the identification of gaps/needs and brokering data access agreements

- Providing a lobbying and representative role, with central funders and other key stakeholders

- And many more….

Survey results - 1

2008 survey of members provided limited information about availability of Census data.

Of 10 organisations:

- 6 distribute Census data
- 4 distribute administrative registers

Survey results - 2

Of 17 responding organisations:

- 76.5% distribute demographic & population data

- 41.2% distribute transport, travel & mobility data

- 35.3% (of 17) distribute housing, land use & planning data

Progress to date

- Discussions with:

  - European funding agencies, to support national archives and promote data sharing

  - DG Research in support of their policy for sharing EU funded SSH research data and related outputs – CESSDA *may* provide a dissemination support service

  - Eurostat and NSI's about better access to official data, including confidential data

Next steps

Upcoming FP7 call (*provisional* information):

- – Time: Summer 2009

- – Scale: Millions of euro

- – Goal: Collaboration between CESSDA and Eurostat/NSI's

- – Possible proposal for pilot project on feasibility of a European SAS – early stages of discussion with ESRC Census co-ordinator

For more information about

the ERI project:

http://www.cessda.org/project/

and about ESDS services:

http://www.esds.ac.uk/

Appendix – links to resources

- IPUMS-I: https://international.ipums.org/international/
- Eurostat surveys: http://epp.eurostat.ec.europa.eu/portal/page?_pageid=1913,47567825,1913_47567846&_dad=portal&_schema=PORTAL
- CESSDA portal: http://www.cessda.org/accessing/catalogue/index.html
- ESDS International: http://www.esds.ac.uk/international
  - Guide to freely available data, *inc. list of NSIs, NGOs*: http://www.esds.ac.uk/international/resources/guide.asp
- UK Census
  - SARs: http://www.ccsr.ac.uk/sars/access/nonacademic/
  - Longitudinal study: http://www.ons.gov.uk/about/who-we-are/our-services/longitudinal-study/using-the-ons-longitudinal-study/access-and-user-support/applying-to-use-data/index.html
  - ONS: http://www.ons.gov.uk/census/index.html
  - Academic access: http://www.census.ac.uk/
- Eurostat aggregate-level data: http://epp.eurostat.ec.europa.eu/portal/page?_pageid=1090,30070682,1090_30298591&_dad=portal&_schema=PORTAL
- UK Neighbourhood Statistics: http://neighbourhood.statistics.gov.uk/dissemination/LeadHome.do;jessionid=ac1f930b30d7d1260e45455945abad49ac4acca3da0a?m=0&s=1238072650781&enc=1&nsjs=true&nsck=true&nssvg=false&nswid=1003
- UK Nomis: https://www.nomisweb.co.uk/