User Group Knowledge Exchange Event
12th July 2013
Scott Orford and Ian Harvey
Wales Institute of Social and Economic Research, Data and Methods (WISERD)

wiserData
- Data key resource for social science research
- Large number of datasets available in Wales/UK **BUT:**
  - Lack of awareness/knowledge about what data exists
  - Discovering data is not straightforward
  - £££ data is under-used
- Awareness and discovery of social science data needs improving (ESRC/HEFCW scoping study – pre WISERD)
  - Existing data needs to be optimised/repurposed/reused (UK Gov Data Strategy, 2009)

WISERD DataPortal
- **Data:**
  - Standards compliant meta-data for search and discovery
    - Quantitative / Qualitative socio-economic data
    - Web mapping services (WMS) feeds
    - API data feeds (NOMIS; ONS)
    - Geographic as well as thematic and temporal metadata searches
- **Portal:**
  - Web-based system for discovering data resources- allows remote 24/7 searching
Mash up of metadata, source data and mapping data

Example Queries

- Text: “Show me all the surveys conducted in Wales that contain questions on health.”
- Spatial: “I need to find out what socio-economic data exists for Cardiff.”
- Spatial: “Show me all the existing data for this (user-defined) area”
- Text/Spatial: “I’m conducting a project on schools in Merthyr Tydfil – I want a printed report of all the relevant data within the Merthyr UA”

WISERD Metadatabase

- WISERD Metadatabase innovations:
  - Metadata for surveys down to question-response level in a fully relational database
  - Semi-automated metadata generation for qualitative research data (keywords and place names)
- Good metadata is the basis for effective:
  - Data discovery
  - Data management
  - Data sharing
  - Data harvesting
  - Data publishing
  - Data archiving
Standards compliant metadata

The WISERD Metadatabase

Surveys:
- Census (1991, 2001)
- National Survey for Wales (2000)
- Shelter Cymru
- Skills Survey (various - 1997-2006)
- Social Change and Economic Life Surveys (1986-1987)
- Welsh Health Surveys (2000-2011)
- Welsh Election Study
- Welsh Referendum Study
- Working in Britain Survey (1996-2001)
- Workplace Employment Relations Surveys (1998-2011)

Stats:
- 129 surveys
- 24 surveys have geo-referenced response statistics
- 14130 questions in DB
- 1043 tables

Qualitative Data: 103 WISERD stakeholder interviews
Administrative Data: 21 PLASC stats tables (2003-10) mapped at UA level. School-level stats mapping planned.
Grey MDB: 10 academic journal articles and PhD theses
Public: No data added yet (future development)
External data sources >>>
DataPortal Core functionality

- Textual database searching
  - Keyword search
  - Advanced text search (under development)

- Geographic/spatial searching
  - Place name + circular distance buffer
  - Drop pin + circular distance buffer
  - Click and drag rectangular search area
  - Draw 'freehand' search area

- Metadata reporting

- Metadata visualisation
  - Mapping
  - Charting
  - Word clouds
Future Developments

• Look to link existing metadata to source data via data feeds
  • NOMIS, ONS
  • INSPIRE & Welsh Government
  • INFUSE (census)
  • data.gov.uk
• Explore methods of publishing/harvesting metadata and data
  – XML based feeds
  – OGC compliant web mapping services (WMS)
• Negotiate and explore direct data access:
  – Secure login (e.g. Shibboleth authentication)
• Explore and manage disclosure risks

Harvesting and publishing external data sources
• Web mapping services (e.g. WG)
• NOMIS
• Welsh Government
• data.gov.uk
• ONS
• ...etc