Supporting Users of Electronic Data Resources in Research and Teaching

Stuart Macdonald  
Edinburgh University  
Data Library

Luis Martinez  
London School of Economics  
Data Library
• What is data?
• Where do you get it from?
• Support Services
• Developments in data storage, dissemination and analysis
• User Experiences
• Conclusions
What is Data?

Some definitions:

a collection of observations or other information related to a particular question or problem

information, most commonly in the form of a series of binary digits, stored on a physical storage medium for manipulation by a computer program

a collection of facts from which conclusions may be drawn

a single piece of information; a fact; especially a piece of information obtained by observation or experiment

information in numerical form that can be digitally transmitted or processed
Data Types

Micro data
- Large-scale government surveys
- Population censuses
- Election and public opinion surveys

Aggregated data
- Macro-level economic indicators

(Geo-)Spatial data
- Census/administrative boundary data
- Satellite imagery
- Base mapping - raster/vector

Financial data
- Company-level data
Micro data

Aggregate data

Spatial data

Financial data

e-Resources in HE FE: changes, challenges and choices

University of Manchester, July 2006
National Statistics

Office of National Statistics (ONS)
- Government department that provides UK statistical services
- ONS is responsible for producing a wide range of key economic and social statistics which are used by policy making

General Register Office for Scotland (GROS)
- responsible for registering vital statistics in Scotland in addition to the Scottish Census

Northern Ireland Statistics and Research Agency

Statistics for Wales – National Assembly for Wales
JISC-funded UK Data Centres

**EDINA** - hosted by Edinburgh University Data Library

**MIMAS** – run by Manchester Computing at the University of Manchester

**UK Data Archive** – based at University of Essex

**Economic and Social Data Service, ESDS** – collaborative and distributed services

**Arts and Humanities Data Services** – distributed services
Free Resources

Large range of statistics and data available online:

- Non-Governmental Organisations
- National/Regional Statistical Institutes
- Government websites (national/local)
- Independent Research Organisations
- Charitable Organisations

Data Discovery Tools:

- Intute Social Sciences – Statistics and Data Section
- Go-Geo!
- Reviewed internet resources e.g.
  - *Edinburgh University Data Library External Sources*
  - *ESDS International’s “Guide to freely available international data resources”*
Commercial Services

There are many commercial data resources/products/services offered at discounted rates to the tertiary education community:

- Datastream, FAME, AMADEUS
- Ordnance Survey Data
- Address Manager/Finder

In addition to licensed commercial software such as:

- ArcGIS/MapInfo (for spatial data analysis)
- SPSS, STATA, SAS (statistical analysis packages for micro-data)
Data Support Services

Institutions provide support for data services in different ways:

- Data Libraries
- University Libraries
- Computing Centres
- Research Offices
- Academic Departments

Data Libraries go beyond the support of national data centres & statistical agencies:

- Act a ‘repository’ of data
- Reference service
- Train users to access and handle data resources
UK Data Libraries

Edinburgh University Data Library - first such service in the UK, 1983

Oxford University Data Library –1988

London School of Economics Data Library –1997

RLab Data Service –1999, providing support to LSE’s research laboratory

Other institutions with ‘Social Statistics’ libraries:

• University of Southampton
• Strathclyde University
Other Support

**DISC-UK** (Data Information Specialist Committee – UK)
A support group formed in 2003 whose aims are as follows:

- Foster understanding between data users and providers
- Raise awareness of the value of data support in Universities
- Share information and resources among local data support staff

**IASSIST** (International Association for Social Science Information Service & Technology)

An international organization of professionals working in and with information technology and data services to support research and teaching in the social sciences
Institutional Repositories

UK Repository Projects:

- StORe – Source-to-Output Repositories
- GRADE - Geospatial Repository for Academic Deposit and Extraction
- R4L – Repository for the Laboratory
- SPECTR Ra
- CLADDIER

Proprietary Repository Systems:

- SDA – Survey Documentation & Analysis
- VDC – Virtual Data Center
- NESSTAR? - Networked Social Science Tools and Resource
Role of data in eScience

eScience

“...is about global collaboration in key areas of big science and the next generation of infrastructure that will enable it” Dr John Taylor

Key words to understand eScience

• Collaboration
• Data
• Research
• Infrastructure
The GRID

The user can look at data from databanks elsewhere and perform analysis using the powerful computers of the GRID.
Reviews of User Experience


• Barriers
  • Problems locating data
  • Lack of awareness of available data
  • Lack of local user support
  • Different service providers, multiple registration systems and interfaces
Locating data

Metadata from national data centres shared across institutions
- National data centres to be 'localised' (Rice et al. 2001)

Metadata standard for Social Sciences Data:
Data Documentation Initiative, DDI version 3.0
Awareness of available data

Data Centres and Libraries coordinated promotion to raise awareness

Training courses, workshops and conferences
- Links to local data support staff/teaching staff

Other channels of communication
- Mailing list
- RSS feeds\Blogs\Wikis

Libraries need someone knowledgeable about data resources and support
- Understanding local needs
- Data sources
- Other support areas (statistical support/research methodology/survey design)
- Support group and associations (IASSIST/DISC-UK)
Local data support

- Locating data and raising awareness
- Troubleshooting
- Hands-on training
- One to one help
- Easing tedious registration
- Streamline ordering
- Manage institutional subscriptions
Easier Access

• Reducing number of interfaces

• **Authentication methods**
  - Global usernames and passwords (ATHENS)
  - Shibboleth

• Data aggregators (beyond 20/20 or Nesstar)

• Grid technologies
Outcomes & Recommendations

- Metadata for all to make data easier to find
- Closer collaboration between local support and national data centres
- Sufficient resources allocated to data support at the institutional level
- Adoption of new technologies for data access and data management