

## Access to detailed social survey micro data

The majority of data held at the UK Data Archive (UKDA) are fully anonymised with certain variables, such as low-level geographies, suppressed to minimise the risk of disclosure.

However, it is recognised that some researchers need access to more detailed data. To facilitate this and to increase the range of social science datasets available for statistical research, the Office for National Statistics (ONS), in collaboration with the UKDA/ESDS, has developed a strategy to provide, through the UKDA, access to data that are detailed, yet still anonymised. Since these data pose a higher risk of disclosure, they have additional special conditions attached to them that take the form of a Special Licence (SL).

The SL specifies the conditions of access, obligations of the researcher, and measures for protecting and respecting the confidentiality of the data. Data are only provided to those researchers who can justify their need for the data and agree to meet the terms of the SL.

*A Guide to Good Practice: Micro Data Handling and Security* has been produced. This

includes guidance on: avoiding inadvertent statistical disclosure; storing and accessing data securely; how to permanently destroy copies of data; and institutional responsibilities, including compliance with security standards.



The first ONS social survey to trial this new access initiative is the Annual Population Survey (APS) 2004 which is now available via the UKDA online catalogue. The APS represents a major new survey of approximately 65,000 achieved household interviews conducted each year. Key topics in the survey include education, employment, health and ethnicity. The APS provides high quality data that can produce reliable estimates at local authority level.

[www.data-archive.ac.uk/findingData/apsTitles.asp](http://www.data-archive.ac.uk/findingData/apsTitles.asp) ■

## New publication from the UKDA and TNA

The UKDA is pleased to announce the publication of *Assessment of UKDA and TNA Compliance with OAIS and METS Standards* resulting from a recent collaboration between the UKDA, The National Archives (TNA) and the Estonian Business Archive.

The publication, and the work on which it is based, was made possible following an award from the JISC under its Institutional Digital Preservation and Asset Management Programme. The award enabled the UKDA and TNA to test their compliance with the Open Archival Information System (OAIS) Reference Model and allowed each organisation to place a recognised 'seal of approval' on established practices and systems. It also enabled the two organisations to compare their preservation practices within a common framework: an



opportunity that was particularly timely because, in January 2005, the UKDA was appointed as a legal place of deposit for TNA.

The award provided the opportunity for the two organisations to explore the value of the Metadata Encoding and Transmission Standard (METS) for a digital archive and to review the application of METS at the UKDA and at TNA. By recording the experiences of, and methodology for, assessing organisational correspondence with the OAIS Reference Model, the report aims to be of use to other organisations that might want to undertake a similar exercise. It can contribute to any future adaptation of the OAIS model to the JISC Information Environment.

[www.data-archive.ac.uk/news/publications.asp](http://www.data-archive.ac.uk/news/publications.asp) ■

## ONS Survey Charter

ONS has recently launched the ONS Survey Charter. Based on the National Statistics Code of Practice, the Charter sets out ONS' commitments to survey respondents and data users.

The Charter explains how ONS aims to meet the needs of government and the wider user community and the steps taken to protect the confidentiality of survey respondents. It details how ONS uphold the quality of their statistics, ensure openness and transparency, and promote widespread access, understanding and use.

[www.statistics.gov.uk/about/ons/survey\\_charter.asp](http://www.statistics.gov.uk/about/ons/survey_charter.asp) ■

# US workshop on data sharing and dissemination

In June Louise Corti was invited as a guest lecturer to run a workshop session on archiving and sharing qualitative data as part of the Inter-university Consortium for Political and Social Research (ICPSR) Summer Program, on *Data Sharing & Dissemination: Making Your Data a Resource for Others*. Data producers were introduced to a range of relevant topics including: designing and structuring variables for ease of analysis; producing longitudinal data files; addressing disclosure limitation; developing effective documentation and the Data Documentation Initiative; creating an effective user guide;

establishing quality control measures for public-use data; qualitative data; analysing data through online analysis systems; developing data sharing plans for restricted-use data; Review Boards; publicising data availability; and preserving data in perpetuity. This was the first time qualitative data had been included by ICPSR in a depositor-oriented training event such as this, pointing to a greater international acceptance of the role of sharing qualitative research data.

[www.icpsr.umich.edu/training/summer/courses/2005-03-40.html](http://www.icpsr.umich.edu/training/summer/courses/2005-03-40.html) ■

## Social Inequality: Using Evidence from Data

On 20 June ESDS hosted a one-day conference on social inequality at the Royal Statistical Society in London, as part of ESRC Social Science Week. The conference highlighted the value of empirical evidence based on ESDS-supported data from large-scale government surveys, longitudinal surveys, international data sources and qualitative data collections. Peter Elias, the ESRC Strategic Advisor for Data Resources chaired the event.

Tania Burchardt, from the ESRC Research Centre for Analysis of Social Exclusion (CASE) at the London School of Economics and Political Science (LSE), opened with a stimulating talk, *Dimensions of Inequality*. Richard Berthoud, from Institute for Social and Economic Research (ISER), University of Essex, presented *The Dynamics of Deprivation* and Roy Carr-Hill, from Centre for Health Economics, University of York, introduced an international perspective by addressing dimensions of inequality using international development

data. Alice Sullivan, from the Institute of Education, spoke on her work, on students' academic attainments, using the National Child Development Study (NCDS).

David Blackaby, from University of Wales, Swansea, described interesting evidence from analysis of the Labour Force Survey (LFS) on Catholic/Protestant Earnings and Employment Differentials in N. Ireland. Julia Johnson, from the Open University, ended on a well-received talk about her qualitative work re-using Peter Townsend's *The Last Refuge*, on continuity and change in residential care for older people.

[www.esds.ac.uk/news/eventdetail.asp?ID=1352](http://www.esds.ac.uk/news/eventdetail.asp?ID=1352)

ESDS has also produced a guide on social inequality, social exclusion and poverty issues, as part of its new series of thematic guides.

[www.esds.ac.uk/support/guides.asp](http://www.esds.ac.uk/support/guides.asp) ■

## E-resources for Research in the Humanities and Social Sciences

The British Academy has consulted national and representative bodies, and individual researchers, and has examined the technical and organisational issues of e-resource provision and access. Its survey and analysis has led to a series of recommendations addressed to institutions and individual researchers.

[www.britac.ac.uk/reports/eresources/index.html](http://www.britac.ac.uk/reports/eresources/index.html) ■

## AHDS History News

### Collection of Historical and Contemporary Census Data

As part of the JISC-funded Collection of Historical and Contemporary Census Data (CHCC), AHDS History has improved access for UK higher and further education users to a complete transcription of the 1881 census enumerators' books for England and

Wales, the Channel Islands and the Isle of Man (data for Scotland are not yet available). The system allows users to request the data by geographical area (registration county, registration district, sub district and parish) and filter by gender and marital status.

Enhanced access is also provided to two major collections of aggregate level demographic data - the Great Britain Historical Database and the Database of Welsh Historical statistics.

[ahds.ac.uk/history/collections/chcc.htm](http://ahds.ac.uk/history/collections/chcc.htm) ■

# Inside the data factory

## Ingest processing

'Ingest processing' describes the work of the UKDA Data and Support Services team to prepare data for secondary use.

Following the acquisitions stage, the deposited dataset undergoes an initial review, to identify any potential problems. For example, data are checked to ensure confidentiality is not compromised and that documentation supporting the data collection is adequate.

Ingest processing is carried out to pre-defined standards, depending on the nature and condition of the dataset and likely secondary usage.

Quantitative data processing ensures that data are accurate and are sufficiently labelled for secondary use; labels may be added/edited according to the assigned processing standard. Accompanying documentation is also scrutinised closely to ensure accuracy and files are combined into a user guide, generally in Adobe Portable Document Format (PDF).

Many qualitative datasets comprise a set of transcribed interviews. During processing, the data are examined to ensure they are complete, that transcription is of high quality, and that confidentiality has been maintained.

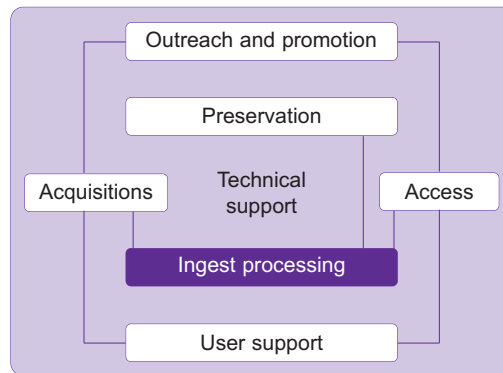
A standardised listing of material is compiled to aid secondary analysis. As for quantitative datasets, a user guide is created from documentation.

For all datasets, a comprehensive catalogue record is created to international metadata standards. A UKDA information file is also produced, which includes details of processing and any further guidance necessary for secondary use.

The dataset is then archived within the UKDA preservation system and released for registered users to order or download.

A Service Level Definition (SLD) is in place to ensure that UKDA ingest processing is completed within 30 working days for quantitative datasets and 35 working days for qualitative deposits. In practice, many datasets are processed very quickly. The efficiency of UKDA acquisitions procedures ensures that most deposits are in excellent condition and the initial review stage uncovers most potential problems. A range of quality-controlled data tools has been developed in-house, which automate elements of ingest processing. Work continues to produce new tools and improve existing ones, so that procedures may become further streamlined and efficient.

Further information about ingest processing can be found on the ESDS web site. ■



## UKDA awards update

### SAFARI

Shibboleth Authentication for Access to the Resource Infrastructures of the UKDA (SAFARI) is a 12-month, JISC-funded project investigating, developing and testing a system that will apply Shibboleth middleware to three resources hosted at the University of Essex. The three resources are: ESDS, the Census Registration Service and the CHCC Collection of Historical Censuses.

Shibboleth provides a way of exchanging information between an organisation and a provider of digital resources, protecting both the security of the data and the privacy of the individual. ■

### StORe

UKDA has been awarded JISC funding for the Source to Output Repositories (StORe) project to provide use cases, guidelines and tools for the integration of data archives with repositories of research publications. ■

## Tools and utilities

A new download area, set up in September 2005, forms part of a UKDA initiative to make its data and metadata expertise freely available to the outside world. The area will contain a growing number of utilities for generating, managing, manipulating and converting social survey data and metadata files.

Tools developed by the UKDA and ESDS will be directly downloadable, complete with accompanying documentation. Links to other similar useful tools made freely available by other organisations will also be provided.

The SPSS to STATA data conversion tool will be the first freely-available means of fully translating data file contents between these two packages. Uniquely, it also generates a log of any information inevitably lost upon conversion (given the two packages have slightly different data handling limits), and a data dictionary/codebook file relevant to both the SPSS and STATA versions of the data.

One of the special features of the conversion utility is the appending of the SPSS user missing value to the beginning of the value label in STATA, making tracking of the conversion very easy.



Kay Eastaugh, Sharon Bolton and Alasdair Crockett (Data and Support Services)

[www.data-archive.ac.uk/support/](http://www.data-archive.ac.uk/support/) ■

## ESDS International's first annual conference

ESDS International will be holding its first annual conference on 8 November 2005 at the Royal Statistical Society in London.

The theme of the conference is the use of international macro data in research and teaching.

The programme will include research papers presented by users of international data on subjects such as international trade and business cycles, the World Bank's illiteracy figures, UN's Millennium Development Goals and global energy-environment-economy modelling. There

will also be presentations from two of the service's data providers - the Organisation for Economic Cooperation and Development (OECD) and the United Nations Industrial Development Organization (UNIDO).

The conference will provide an opportunity for members of the academic community using the international databanks in their research and teaching, data librarians and producers of international data resources to meet and share ideas.

[www.esds.ac.uk/international/news/conf2005.asp](http://www.esds.ac.uk/international/news/conf2005.asp) ■

## Online qualitative data: using the right standards

QUADS - the ESRC Qualitative Demonstrator Scheme, co-ordinated by ESDS Qualidata is holding its first workshop on Online Qualitative Data Resources: Best Practice in Metadata Creation and Web Standards, on 15 November 2005 at Centre Point, London.

The day will discuss the fundamentals of presenting qualitative data resources on the web and for longer-term sharing and archiving, with the focus on promoting current standards and gaining support for emerging standards.



[www.esds.ac.uk/qualidata/news/eventdetail.asp?ID=1441](http://www.esds.ac.uk/qualidata/news/eventdetail.asp?ID=1441) ■

## Forthcoming Events

**5 October 2005:** Survey Link Scheme Workshop with collaboration from ONS Labour Force Survey, University of Surrey

**12, 20, 21, 26 October 2005:** AHDS Digitisation Workshops, Oxford, London, York

**14 October 2005:** Developing Labour Market Intelligence - a seminar by the Labour Market Statistics User Group, Royal Statistical Society, London

**18 October 2005:** Introductory workshop on a social capital theme, ONS, Drummond Gate, London

**3-4 November 2005:** Secondary Analysis in Qualitative Research. Utopia and Perspectives, Grenoble, France

**8 November 2005:** Smaller Households, Smaller Families: implications for policy, planning and support, University of Edinburgh

**8 November 2005:** Race and social research: issues of methodology, ethics and practice, York

**8 November 2005:** ESDS International Annual Conference 2005, Royal Statistical Society, London

**11 November 2005:** The Analysis of Qualitative Longitudinal Data for Policy, Friends Meeting House, Euston Road, London

**11 November 2005:** Survey Link Scheme Workshop with collaboration from NATCEN National Travel Survey, University of Manchester

**12 November 2005:** Association for History and Computing UK 2005 Conference, University of Cambridge

**15 November 2005:** Online Qualitative Data Resources: Best Practice in Metadata Creation and Web Standards, Centre Point, London

**16-17 November 2005:** Census: present and future, University of Leicester

**6 December 2005:** Labour Force Survey User Group Meeting, Royal Statistical Society, London

**15 December 2005:** NCDS/BCS70 User Group Meeting: Education, skills and training in NCDS and BCS70, Institute of Education, London

Further details of events and training courses can be found on the ESDS web site.

### News Byte

During August, 2664 datasets were downloaded from the web site.

Editorial group enquiries  
Email: [publicity@esds.ac.uk](mailto:publicity@esds.ac.uk)  
Tel: +44 (0)1206 872319/873159  
Fax: +44 (0)1206 872003

General ESDS enquiries  
Email: [help@esds.ac.uk](mailto:help@esds.ac.uk)  
Tel: +44 (0)1206 872143  
Fax: +44 (0)1206 872003

Acquisitions  
Email: [acquisitions@esds.ac.uk](mailto:acquisitions@esds.ac.uk)  
Tel: +44 (0)1206 872572/872974  
Fax: +44 (0)1206 872003

ESDS Government  
Email: [govsurveys@esds.ac.uk](mailto:govsurveys@esds.ac.uk)  
Tel: +44 (0)161 275 1980  
Fax: +44 (0)161 275 4722

ESDS International  
Email: [international@esds.ac.uk](mailto:international@esds.ac.uk)  
Tel: +44 (0)161 275 6109  
Fax: +44 (0)161 275 6071

ESDS Longitudinal  
Email: [longitudinal@esds.ac.uk](mailto:longitudinal@esds.ac.uk)  
Tel: +44 (0)1206 872143  
Fax: +44 (0)1206 872003

ESDS Qualidata  
Email: [qualidata@esds.ac.uk](mailto:qualidata@esds.ac.uk)  
Tel: +44 (0)1206 873058  
Fax: +44 (0)1206 872003

AHDS History  
Email: [info@history.ahds.ac.uk](mailto:info@history.ahds.ac.uk)  
Tel: +44 (0)1206 872326  
Fax: +44 (0)1206 872003

Economic and Social Data Service  
UK Data Archive  
University of Essex  
Wivenhoe Park  
Colchester  
Essex CO4 3SQ, UK