

# UK DAtabytes

The Newsletter of the UK Data Archive

No. 8

## Inside the new Economic and Social Data Service

The ESDS is a new national data archiving and dissemination service which came into operation in January 2003. The service has initially secured funding for just under five years, and is a jointly-funded initiative sponsored by the Economic and Social Research Council (ESRC) and the Joint Information Systems Committee (JISC).

The ESDS has been established as a distributed service, based on a collaboration between four key centres of expertise: the UK Data Archive (UKDA), the Institute for Social and Economic Research (ISER), both based at the University of Essex; Manchester Information and Associated Services (MIMAS) and the Cathie Marsh Centre for Census and Survey Research (CCSR), both located at the University of Manchester.

The ESDS comprises six distinct yet inter-related functions.

### Management and Co-ordination

The overall direction and management for the ESDS will be the responsibility of the UKDA, under the leadership of Kevin Schürer. Through this function, the UKDA will provide consistency and standards across the service. It will have a broad strategic role, relating to a variety of stakeholders concerned with the supply, funding and use of social science data, and will create a coherent publicity, promotion and outreach strategy for the whole service. The key deliverables include:

- a *universal portal* for social science data resources, providing a flexible resource discovery tool for key ESRC and JISC-funded social science data services and other relevant resources;
- a *central registration* service operating across the ESDS enabling users to move seamlessly between different online resources;
- a central 'first stop' *help desk* service;
- a new *collections development* policy, incorporating a systematic review of all potential data acquisitions in line with users' needs;
- a co-ordinated *outreach, training and promotional* strategy.



ESDS UKDA management team (l - r) Louise Corti, Melanie Wright, Kevin Schürer, Diane Geraci.

### Data Archiving, Preservation and Dissemination Service

Hosted at the University of Essex and undertaken by the UKDA, this service represents the core of the ESDS and incorporates most of the activities that are currently performed by the UKDA. Focusing on the central activities of data acquisition, processing, preservation and dissemination, the key services to be provided are:

- a *collections development strategy* working to identify and acquire a range of data from across the social sciences and related disciplinary areas as appropriate - whilst most of the data will be in numeric and textual form, other formats, such as image and audio files, may also be considered;
- *ingest activities* including data and documentation processing, validation, consistency and cross-comparison checking, together with the checking of licensing arrangements, copyright and access conditions and conversion of data to differing formats appropriate for preservation and dissemination;
- *metadata creation* including online user guides, bibliographic catalogue records and assigned index terms to aid resource discovery;
- a *data dissemination* service via an instant web-based download facility and the writing of bespoke CD-ROM/DVDs;
- *long-term preservation* for all datasets acquired across ESDS, monitored to account for technological shifts, changes in perceived best practice and the nature of ESDS holdings.

Supported by:



University of Essex

E · S · R · C  
ECONOMIC  
& SOCIAL  
RESEARCH  
COUNCIL

JISC

Special Edition

Serving and preserving digital resources for research and learning

## Large-scale Government Data Service

The large-scale government surveys, such as the General Household Survey, the Labour Force Survey and the Family Expenditure Survey are key data resources for social science researchers. The service will promote a support and training programme designed to generate a significant increase in the use of these datasets among both potential novice users and more sophisticated data analysts.



CCSR team (l - r) Jo Wathan, Sam Smith, Vanessa Higgins, Angela Dale.

The service will be led by the CCSR at Manchester under the direction of Angela Dale, and will:

- create a number of *value-added products* providing easy routes through the complex documentation of the government surveys, including methodological guides, key sets of derived variables, and contextual information for the historical time periods relevant to the surveys;
- provide support for the three main *statistical packages*, SPSS, STATA and SAS;
- provide *specialist advice*, through a help desk service, a moderated discussion list and a web-based FAQ site, on data management, statistical advice on analysis, model building and interpretation of results for SPSS and STATA;
- run *training courses* to raise awareness of government datasets and their research potential on key topics of interest, on specific statistical packages, and on methods of statistical analysis, covering a range of levels;
- provide topic-related online *course materials* and a range of *teaching datasets* developed and updated in consultation with teachers;
- work closely with ONS and other data producers to ensure that the complementarity of interest between producers, commissioning government departments and academic users is maximised.

The acquisition, processing and delivery of the large-scale government datasets under this service will be undertaken by the UKDA. In

addition to studies being made available via the ESDS web-based download system, the large-scale government surveys supported by this service will also be mounted in the NESSTAR system for online browsing and visualisation of the data.

## International Data Service

This service will provide access to, and support for, a range of international datasets – both macro and micro sources. The central aims of the service are to promote and facilitate increased and more effective use of international datasets in research, learning and teaching across a range of disciplines. The service will be run jointly by MIMAS at Manchester, under the leadership of Keith Cole, and the UKDA at Essex and will concentrate on:

- developing a *responsive acquisitions strategy* that negotiates key international datasets for UK academic access;
- providing *web-based access* to regularly updated macro-economic time series datasets, provided by organisations such as the OECD and the IMF, using Beyond 20/20 software;
- helping users to locate and acquire international *micro-level datasets*, such as Eurobarometer or ISSP, from other social science data archives via a series of reciprocal agreements;
- developing key time series for *alternative classifications* of countries, such as highly indebted countries;



MIMAS team (l - r) Celia Russell, Jim Schumm, Gindo Tampubolon, Keith Cole.

- using CommonGIS software to build an *interactive web-based data exploration* and visualisation interface to a set of key international statistics;
- providing *specialist support* on international datasets via a dedicated telephone and email help desk, and user forums;
- developing *web-based materials* in the form of introductory guides covering the structure and content of international data together with more advanced user guides covering key methodological issues and concepts, such as harmonisation and cross-country comparability;
- running a *programme of courses* to raise awareness of international datasets and their research potential and on data analysis.

“... raise awareness of government datasets...”

“... facilitate increased and more effective use of international datasets...”

The service will build upon its existing network of contacts to develop productive relationships with key data providers and international social science data archives.

## Qualitative Data Service

This specialist service will provide access to, and support for, a range of qualitative datasets, and will be hosted by the UKDA, with the team being led by Louise Corti. The work builds on Qualidata expertise and international reputation in this area, developed over the past eight years. A six-fold strategy of data enhancement will be developed comprising:

- the creation of *web-based samplers* aiming to provide 'edited highlights' of key qualitative materials to illustrate the potential of the collection for research and teaching;
- the creation of *thematic resources* whereby interviews relating to a particular theme and time period will be combined into a single resource, for example, crime and social order in the late-twentieth century;
- *value-added processing*, ensuring that interview data are fully anonymised, are in an appropriate digital format, have speakers' tags assigned and have enhanced finding aids comprising dedicated online user guides, and associated web pages;
- *web delivery* of marked-up primary data, such as interview transcripts, using XML standards and tools to facilitate rapid and flexible retrieval of information – as demonstrated in the Edwardians Online project representing the first phase of this plan of work;
- *enhanced access* to key qualitative collections held elsewhere, in partnership with the hosting archive, to facilitate use in research and teaching;
- a *video-archive demonstrator*, that aims to investigate the use of video methods, focusing on methodological, ethical, technical and analytical questions.

The projects selected for data enhancement will be identified in consultation with key stakeholders, the wider research and teaching communities and the Qualidata Advisory Committee.

The publishing of enhanced qualitative data resources within web-based systems will be supplemented by mounting all newly-acquired core qualitative data on the UKDA's web-based direct download service. Previously acquired electronic data will be mounted on the download service according to user demand. Data will ordinarily be offered in word-processed or plain text formats rather than in a specific Computer Assisted Qualitative Data Analysis Software (CAQDAS) format, but the service will also meet any significant demand for software specific formats. One of the service's early priorities is to define and encourage the uptake of a software-independent data format both for longer-term preservation purposes, and for transporting coded data between CAQDAS packages.

Great emphasis will be placed on user support, and a dedicated help desk facility will be established. Tailored user guides, theme-based web pages and FAQs will be produced. User events and training

days will be organised in line with identified user needs. Workshops will include generalist introductory sessions and more focused meetings on detailed areas of research interest and methodology. These will be supplemented by 'data confrontation' workshops aiming to enhance the methodological and substantive understanding, and secondary analytical potential, of archived qualitative data sources.

Advice and support will also be provided for creators and depositors of qualitative data to ensure that the acquired data are fully supported by adequate documentation, and that confidentiality, data protection and copyright issues have been addressed from the beginning of the research.

## Longitudinal Data Service

The work of the Longitudinal Data Service will be undertaken jointly by the UKDA and the UK Longitudinal Studies Centre (ULSC), at the University of Essex. Nick Buck will be leading the team from ULSC, based at the Institute for Social and Economic Research (ISER). The service will support a range of longitudinal data collections used by UK social scientists and will encourage linkage with other datasets not directly supported by ESDS, such as the ONS Longitudinal Study.



Nick Buck, Associate Director ISER/Director ULSC.

This service will provide value-added support for the following data collections: the British Cohort Study (BCS70), the British Household Panel Survey (BHPS), the Millennium Cohort Study (MCS) and the National Child Development Survey (NCDS). It will work closely with the ESRC's national longitudinal strategy, managed by the ULSC, in conjunction with the ESRC National Strategy Committee. With reference to this existing body, together with the acquisition strategy set by the Management and Co-ordination function of the ESDS, this service will further identify new longitudinal data collections suitable for inclusion under this service. Possible candidates would include the English Longitudinal Study of Ageing (ELSA), the Avon Longitudinal Study of Parents and Children (ALSPAC), the 1946 Birth Cohort and the planned UK New Longitudinal Survey of Young People in England.

Value-added *data enhancements* will be undertaken aimed at aiding the usability of the

"... enhanced access to key qualitative collections..."

"... support a range of longitudinal data collections..."

## Social Data Service

*Serving and preserving digital resources for research and learning*

available datasets including:

- providing *harmonised data* standards across all waves or sweeps, and assurance of the integrity of case identifiers;
- improving information on *weighting*, the presentation of *missing data* and other statistical adjustments, and additional derived variables – the ULSC is already undertaking this work for the NCDS, BCS70 and BHPS;
- developing an *online system* for longitudinal data that will allow the visualisation of documentation, and a front-end to the system that will further facilitate the sub-setting and data matching required to produce a range of possible final target structures;
- developing means of allowing researchers to make use of *confidential data*, especially geo-coded data about members of longitudinal studies, without prejudicing the confidentiality of those respondents.

A three-fold strategy for dissemination of longitudinal data collections will be developed comprising:

- *instant web-based* downloadable data files;
- a *data management tool*;
- the creation of *online sampler files* to facilitate the primary investigation of the data and evaluation for research and teaching.

The UKDA will create a series of easy-to-navigate sampler datafiles for the BHPS, NCDS and BCS70 that will either form time-slices through the data in order that users can explore online the content and extent of the data collection, or present derived variables such as poverty indicators. These samplers will be specially prepared and mounted online within the NESSTAR system. To accompany such data samplers downloadable 'workbooks' will be produced giving simple examples of exploratory analysis and longitudinal regression. These teaching samplers will be developed with user consultation defining the topic areas.

The service will also establish a specialist longitudinal data help desk to provide advice on a variety of technical and non-technical aspects of

data analysis. This service will link up with further user support provided at the Centre for Longitudinal Studies at the Institute of Education.

In order to promote the exploitation of longitudinal data in new and emerging user communities, a series of meetings and workshops aimed towards novice users will be held. More advanced users will be supported by a number of focused workshops, training and 'data confrontation' sessions. A programme of training will be run in conjunction with the Essex Summer School Programme. All training materials produced in conjunction with workshops and training events will be made available via the web site. A further facility will be a searchable online directory of researchers' interests, an online catalogue providing details of other longitudinal data collections, and information about the development of the ESRC's National Strategy for Longitudinal Studies.

---

## ESDS launch event in London

During the afternoon of 30 June 2003 the ESDS will be hosting its inaugural launch, at Regent's College in central London. Ian Diamond (Chief Executive ESRC), Len Cook (the National Statistician) and ESDS staff will be speaking about the range of exciting and forward-looking opportunities that the new ESRC/JISC data service can offer to social scientists and beyond. As places are limited, those wishing to attend should register their interest by contacting [archive@essex.ac.uk](mailto:archive@essex.ac.uk) by 1 May at the latest. There will be no charge for the day.

---

## Next steps

Further information about the service, along with details of the ESDS user survey, are available at [www.data-archive.ac.uk/esds/introduction.asp](http://www.data-archive.ac.uk/esds/introduction.asp)

These news pages will eventually be incorporated into a new ESDS web site.

---

## Forthcoming Events and Announcements

**4 March 2003** JISC Conference 2003, Birmingham.

**2-3 April 2003** National Child Development Study (NCDS) and the 1970 British Cohort Study (BCS70) Workshop, London.

**4 April 2003** Opportunities for complementarity between longitudinal studies, Royal Statistical Society, London.

**4-5 April 2003** Institute for Fiscal Studies (IFS) Eleventh Residential Conference, Oxford.

**7-9 April 2003** Knowledge and Economic and Social Change: New Challenges to Innovation Studies, Manchester.

**11-13 April 2003** British Sociological Association (BSA) Annual Conference 2003, University of York.

**16 May 2003** BURISA 2003 Conference, London.

**27-30 May 2003** IASSIST 2003 Conference, Ottawa, Canada.

Further details may be found at: [www.data-archive.ac.uk/home/forthcomingEvents.asp](http://www.data-archive.ac.uk/home/forthcomingEvents.asp)

Details of training courses may be found at: [www.data-archive.ac.uk/usingData/training.asp](http://www.data-archive.ac.uk/usingData/training.asp)

Answers to the Christmas competition: A. Ian Diamond, B. Tony Blair, C. Kevin Schürer

---

## Special Edition

*Serving and preserving digital resources for research and learning*

**Editorial Group Enquiries:**  
Email: [archive-pub@essex.ac.uk](mailto:archive-pub@essex.ac.uk)  
Tel: +44 (0) 1206 872319

**General Archive Enquiries:**  
Email: [archive@essex.ac.uk](mailto:archive@essex.ac.uk)  
Tel: +44 (0) 1206 872001  
Fax: +44 (0) 1206 872003

**User Services:**  
Email: [archive-userservices@essex.ac.uk](mailto:archive-userservices@essex.ac.uk)  
Tel: +44 (0) 1206 872017/ 873574

**Acquisitions:**  
Email: [archive-acq@essex.ac.uk](mailto:archive-acq@essex.ac.uk)  
Tel: +44 (0) 1206 872572

**History Data Service,  
The AHDS Centre for History:**  
Email: [hds@essex.ac.uk](mailto:hds@essex.ac.uk)  
Tel: +44 (0) 1206 872326

**Qualidata:**  
Email: [quali@essex.ac.uk](mailto:quali@essex.ac.uk)  
Tel: +44 (0) 1206 873058

© UK Data Archive 2002

Printed by Alphaprint (Colchester) Ltd

**UK Data Archive**  
University of Essex  
Wivenhoe Park  
Colchester  
Essex CO4 3SQ  
UK

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