

# UK DAtabytes

The Newsletter of the UK Data Archive

No. 7

## 'Horizons' workshop on preserving and sharing medical research data

On 8 October, the Medical Research Council (MRC) sponsored a 'Horizons' workshop as part of its Population Data Archiving and Access Project. The workshop explored how data are, and are likely to be, generated, exchanged and combined in the future as science, information and computing technologies develop.

Using exemplars from diverse areas of science, the workshop aimed to:

- identify how different kinds of research data and use call for different preservation and data sharing strategies
- understand how data preservation can promote collaborative data sharing, mining, comparison and integration
- identify potential challenges to, and solutions for, preservation and sharing population data
- develop principles for an MRC policy on preserving and sharing population-based datasets

The workshop was also an opportunity for the UKDA to present the findings of its in-depth case studies on data management, preservation, and sharing of key MRC population-based research data. UKDA staff visited seven MRC research sites and conducted in-depth, semi-structured interviews with principal investigators, project managers, data managers, statisticians and laboratory managers in an attempt to uncover the situation 'on the ground' in terms of attitudes and practices in data management, data sharing, and data preservation. The key findings and recommendations from the substantial consultants' report, produced by the UKDA, were presented at the workshop.

The workshop was structured as a morning plenary with speakers from various areas of e-science, followed by afternoon break-out groups which discussed certain key challenges facing the MRC when developing a data preservation and data sharing policy. These break-out sessions included:

- ownership and custodianship: responsibilities
- infrastructure models
- making choices about what to archive
- confidentiality, consent and access

The workshop concluded with a final report-back session.

Throughout the day, speakers from the fields of genomics, imaging, and informatics demonstrated how information and computing technologies are driving new ways of using data



Horizons organising team (l-r) John Sam, Peter Dukes (project manager), Yanine Jairazbhoy, Sanjiv Chudasama, Clare Davies

and 'doing science'. Information sharing, integration and interoperability pose significant cultural, technical and funding challenges. The key areas emphasised were the importance of metadata and associated standards; the necessity of close collaboration between data specialists, data creators and users in the development of data sharing and preservation standards, tools and infrastructure. Furthermore guidance from the MRC is needed on rights management, data custodianship, consent for future 'new use', quality assurance, and criteria for prioritising resources for investment in preservation and sharing.

### MADIERA project

The UKDA is pleased to announce that we have, together with European partners, secured EU funding to further develop work carried out under the Language Independent Metadata Browsing of European Resources (LIMBER) project.

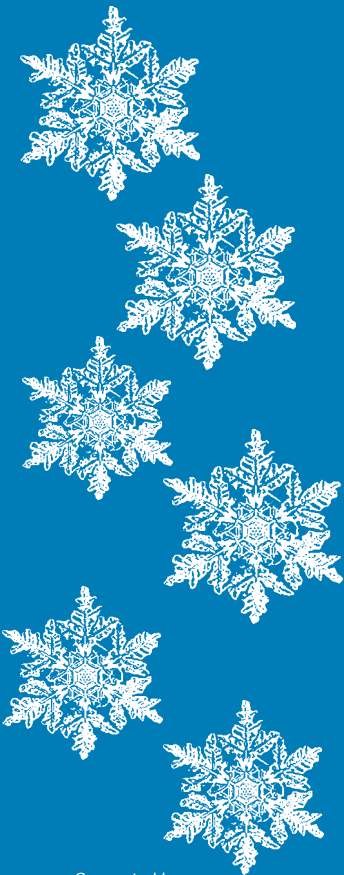
The EU IHP-Programme - 5th Framework is funding the 3-year project, Multilingual Access to Data Infrastructures of the EU Research Area (MADIERA) which will be coordinated by the Norwegian Social Science Data Archives (NSD). The UKDA is one of seven partners of European national data archives.

The overall goal of MADIERA is to develop an effective infrastructure for the European social science community by integrating data with other tools, resources and products of the research process. By the end of the project a fully operational web-based infrastructure will be in place, populated with a variety of data and resources from a selection of providers. The aim is to provide a common integrated interface to the resources of the majority of the existing 20-plus social science data archives in Europe, including several newly-established archives in the candidate countries.



Christmas  
Competition  
Win a bottle of  
champagne

See back page



Supported by:



University of Essex

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

JISC



Seasons Greetings to all our Readers



## Count down to launch of ESDS

In January 2003, the new Economic and Social Data Service (ESDS) will come into being. The service has committed funding from ESRC and JISC for a period of just under five years.

ESDS will be a distributed service with collaboration between four sites and will be managed and co-ordinated by the UKDA team at the University of Essex. Professor Kevin Schürer has been appointed Director, and will be supported by a team of Associate Directors based at the University of Essex and the University of Manchester. The Large-scale Government Data function will be hosted at the Cathie Marsh Centre for Census and Survey Research (CCSR), lead by Angela Dale, with support from the UKDA. The International Data Service will be hosted at Manchester Information and Associated Services (MIMAS) under the leadership of Keith Cole, again supported by the UKDA. The UKDA will have responsibility for the Qualitative Data Service and the core Data Archiving and Dissemination Service, and will share responsibility for the operation of the new Longitudinal Data Service with the UK Longitudinal Studies Centre (ULSC) at the Institute for Social and Economic Research (ISER). The Essex management team will consist of Nick Buck, Louise Corti, Diane Geraci and Melanie Wright.

The overall aim of ESDS is to promote wider and more informed use of data for research and teaching in the social sciences and to ensure that these data continue to be available over time. To achieve this aim ESDS:

- will develop a clear, co-ordinated acquisitions strategy in conjunction with the ESRC and JISC
- develop and promote common standards and agreed strategies for data preparation, processing, documentation and preservation
- integrate data services through the development of a universal single point of access and simple registration procedures and common conditions of use
- enhance underlying data sources through the development of value-added services around key datasets and series
- promote increased use of datasets through both general and specialist user support activities
- develop interface and analysis tools appropriate to differing levels of expertise amongst end users
- respond to key technical developments in the broader information environment
- engage on a regular basis with a wide range of stakeholders, including data suppliers, data funders and end users

The UKDA will be responsible for providing coherence across the ESDS through integrating the activities of the data archiving and dissemination function and the specialist data services. ESDS will have an Advisory Committee, chaired by John Pullinger from the Office for National Statistics (ONS), to advise on the general direction and strategy of the service. This committee will be a representative mix of all stakeholders. Outside of this committee will be a broader network of stakeholders concerned with the supply, funding and use of social science

data. A number of new user forums will be established.

Work is already underway to discuss issues of: governance; branding and promotion; centralised registration; web portal design; data enhancements; and user support and training.

The next issue of UK DATabytes will be devoted to detailing the new service.

## QUALIDATA News

### Discussing enhanced access at Edinburgh conference

In September Louise Corti and Emma Barker attended the Digital Resources for the Humanities (DRH) conference in Edinburgh to present a paper on 'Enhancing Access to Qualitative Data: Edwardians Online'. The DRH is a series of annual conferences aiming to bring together creators, users, distributors, and custodians of digital resources in the humanities. The UKDA paper formed part of the session on 'Resource Creation: Early Books and Manuscripts' which focused attention on the practical and methodological issues surrounding digitising and marking-up text documents.

The UKDA talk introduced the Edwardians Online project and highlighted the need for the development of a comprehensive and inclusive data model that can enable sophisticated online searching and information retrieval of qualitative data. Participants were interested in the principle of, and methodology behind, preserving primary researchers' analytical coding, in addition to the structural features of the dataset. Selected papers from the DRH conferences are usually published in a number of volumes by the Office for Humanities Communication. For further details of the conferences and publications see: [www.drh.org.uk](http://www.drh.org.uk)

## JISC Collection and Service Description Registry

In pursuit of its Information Environment strategy, JISC are compiling a registry of collections and services that are offered across its portfolio of resources. The UKDA will be contributing a number of collection-level records to highlight its major collections, together with a description of the type of service by which these collections can be accessed.

In the context of this exercise, the services of interest are those network services which provide access to collections of digital metadata records. These may be physical, e.g. a book in a library; or digital, e.g. a web site; or a record in a subject gateway database. Network services can be those that provide structured access to metadata records for use by software applications, such as Z39.50 or Open Archives Initiative (OAI) protocols; or those which present metadata in a form intended for a human reader, such as an HTML-based web site.

The UKDA supports the idea of collection description for helping promote its 4000-plus study-based holdings to a wider audience. For the UKDA, a 'collection' can represent, for example, a major government series, such as the General Household Survey; a grouping of datasets with a particular geographic or temporal focus; or datasets classified by data type, such as survey, census, administrative, or qualitative.

*"... to promote wider and more informed use of data for research and teaching"*

*"enhance underlying data sources through the development of value-added services"*



*Seasons Greetings to all our Readers*



## Goodbye to BIRON

As reported in our last newsletter, access to our online catalogue via BIRON will be discontinued after December 2002. The replacement UKDA catalogue search engine has been available on our web site for some time.

The UKDA is sad to be saying goodbye to BIRON after all this time, but advances in technology enables us to provide users with much faster, and more flexible, access to the catalogue.

The UKDA continues to add extra functionality to the new catalogue search engine and still invites feedback.

Email your comments to:  
ukda-feedback@essex.ac.uk

## Award success for UKDA staff

Dr Alasdair Crockett, Data Services Manager at the UKDA, and Dr David Voas of the Department of Sociological Studies at the University of Sheffield have recently been awarded a research grant by the ESRC. The project, entitled *The British Household Panel Study and key issues in religious change*, will run from April 2003 until November 2004. The research will contribute to the knowledge and understanding of religious (and by extension, ethnic) identification through testing and developing theory based on empirical investigation of: the inter-generational transmission of religious identity and activity; the influence of partnership and spousal affiliation on self-classification and religious practice; the impact of various life stages on religious involvement; and the timing of period or cohort effects related to secularisation in Britain.

## Preservation pursuits

UKDA staff attended two important preservation events in October. The Electronic Resource Preservation and Access Network (ERPANET) workshop on 'XML as a Preservation Strategy' was held in Urbino, a well-preserved Renaissance hill town in Italy.

Workshop highlights included an XML prototype for archiving email with associated metadata developed in the Netherlands, a relational database archiving project using SQL-99 and XML at the Swiss Federal Archives, a 'persistent archives' collaborative project between the U.S. National Archives and Records Administration and the San Diego Supercomputer Center, and a European initiative for a broad-based, web archiving project. Presentations from the ERPANET workshop are available at: [www.erpanet.org/](http://www.erpanet.org/) [www/products/urbino/urbino.htm](http://www/products/urbino/urbino.htm)

The Digital Preservation Coalition (DPC) Forum in London focused on two timely themes: 'Preservation of E-Learning Materials' and 'Cost Models for Digital Preservation'. Learning technology standards and digital repository interoperability were promoted alongside creating a digital repository system for research and teaching outputs at a university. Practical cost models for digital preservation based on the Open Archival Information System (OAIS) and digital object life cycle models were presented. Attendees were encouraged to think about the true costs associated with this critical function.

A full report is available from the DPC web site: [www.dpconline.org/graphics/events/elearning.html](http://www.dpconline.org/graphics/events/elearning.html)

## CESSDA Expert Seminar in Romania

This year's Council of European Social Science Data Archives (CESSDA) Expert Seminar was organized by the Romania Data Archive (RODA) and the Zentralarchiv für Empirische Sozialforschung (ZA) and held together with the East European Data Archive Network (EDAN) DDI Workshop.

The seminar provides an opportunity for members of European data archives to share their knowledge and technical expertise. By deciding to hold the 2002 meeting in Bucharest, CESSDA was also able to support EDAN's first DDI training initiative.

In attendance from the UK were Ken Miller and Emma Barker of the UKDA, and Massimiliano Gerardi of Nesstar Ltd. Presentations and hands-on sessions covered a wide range of topics: the DDI and XML codebook production; XML



Delegates from the CESSDA expert seminar

markup of qualitative data; metadata; Nesstar publishing; data access and access control; the MADIERA project; controlled vocabularies and thesauri. A session on CESSDA cross-border data transfer agreements was also productive.

*"... advances in technology enables us to provide users with much faster, and more flexible, access..."*

*"... an opportunity for members of European data archives to share their knowledge"*



*Seasons Greetings to all our Readers*



## Forthcoming Events and Announcements

**17 December 2002** Royal Statistical Society (RSS) Social Statistics Section Meeting: Impact of the new British Crime Survey methodology, RSS, London.

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

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

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

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

**3-5 July 2003** British Household Panel Survey Research Conference, Colchester, Essex.

**3-5 July 2003** Conference of the European Panel Users Network, Colchester, Essex.

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)

**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:**  
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

## History Data Service News

### A Place in History: A Guide to Using GIS in Historical Research

The HDS is pleased to announce the web publication of 'A Place in History', a new guide to using Geographical Information Systems (GIS) in historical research. It will also be published by Oxbow Books next year.

The guide provides a broad sweep of GIS knowledge relevant to historians without assuming prior knowledge. It includes case studies from a variety of historical projects that have used GIS and an extensive reading list of GIS texts relevant to historians.

This guide describes how to create GIS databases and how to use GIS to perform historical research. Its aims are to:

- define GIS and outline how it can be used in historical research
- evaluate the way GIS models the world
- describe how to get data into a GIS
- demonstrate the basic operations that GIS offers to explore a database

- review how time is handled in GIS
- explain how GIS can be used for simple mapping and more advanced forms of visualisation
- discuss quantitative data analysis within GIS
- illustrate the use of GIS for qualitative analysis
- summarise documenting and preserving GIS datasets

It has been commissioned by the HDS as part of the Arts and Humanities Data Service publication series Guides to Good Practice in the Creation and Use of Digital Resources. The series aims to provide guidance about applying recognised good practice and standards to the creation and use of digital resources in the arts and humanities.

The web-based guide is available at: [hds.essex.ac.uk/g2gp/gis/index.asp](http://hds.essex.ac.uk/g2gp/gis/index.asp)



## Christmas competition

### Win a bottle of champagne!



A



B



C

Name the three faces hiding behind the festive disguise. The first, randomly chosen, correct answer will win a bottle of champagne, compliments of the UK Data Archive. The winner, together with the correct answers, will be announced in the next issue of UK Databytes.

Answers in an email to [archive-pub@essex.ac.uk](mailto:archive-pub@essex.ac.uk) by 21 January 2003.



Seasons Greetings to all our Readers



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

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