# UK Data Archive

## ANNUAL REPORT 2001/2002

## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>2</td>
</tr>
<tr>
<td>Services to Depositors</td>
<td>4</td>
</tr>
<tr>
<td>Acquisition</td>
<td>4</td>
</tr>
<tr>
<td>Validation &amp; Conversion</td>
<td>5</td>
</tr>
<tr>
<td>Metadata Creation</td>
<td>5</td>
</tr>
<tr>
<td>Preservation &amp; Collections Management</td>
<td>5</td>
</tr>
<tr>
<td>Services to Users</td>
<td>7</td>
</tr>
<tr>
<td>Dissemination</td>
<td>7</td>
</tr>
<tr>
<td>Resource Discovery</td>
<td>8</td>
</tr>
<tr>
<td>Support</td>
<td>9</td>
</tr>
<tr>
<td>Promotion</td>
<td>9</td>
</tr>
<tr>
<td>Research and Development</td>
<td>11</td>
</tr>
<tr>
<td>Census Registration Service (CRS)</td>
<td>11</td>
</tr>
<tr>
<td>CHCC Project</td>
<td>12</td>
</tr>
<tr>
<td>MRC Data Archiving &amp; Access Project</td>
<td>12</td>
</tr>
<tr>
<td>FASTER</td>
<td>13</td>
</tr>
<tr>
<td>COSMOS</td>
<td>13</td>
</tr>
<tr>
<td>MetaNet</td>
<td>13</td>
</tr>
<tr>
<td>AMRADS</td>
<td>13</td>
</tr>
<tr>
<td>Geo-Spatial Projects</td>
<td>13</td>
</tr>
<tr>
<td>MADIERA</td>
<td>14</td>
</tr>
<tr>
<td>MetaDater</td>
<td>14</td>
</tr>
<tr>
<td>Archives Hub</td>
<td>14</td>
</tr>
<tr>
<td>Collaboration with the RSS</td>
<td>14</td>
</tr>
<tr>
<td>Resource Management</td>
<td>16</td>
</tr>
<tr>
<td>Quality Management</td>
<td>16</td>
</tr>
<tr>
<td>Project Management</td>
<td>16</td>
</tr>
<tr>
<td>Planning &amp; Resource Management</td>
<td>16</td>
</tr>
<tr>
<td>Financial Summary</td>
<td>17</td>
</tr>
<tr>
<td>Qualitative Data Service</td>
<td>18</td>
</tr>
<tr>
<td>New Web Site</td>
<td>18</td>
</tr>
<tr>
<td>Acquisition &amp; Depositing Service</td>
<td>18</td>
</tr>
<tr>
<td>Services to Users</td>
<td>19</td>
</tr>
<tr>
<td>Promotion</td>
<td>21</td>
</tr>
<tr>
<td>History Data Service</td>
<td>23</td>
</tr>
<tr>
<td>Highlights of the Year</td>
<td>23</td>
</tr>
<tr>
<td>Resource Delivery</td>
<td>23</td>
</tr>
<tr>
<td>Training Workshops</td>
<td>23</td>
</tr>
<tr>
<td>Web Site</td>
<td>23</td>
</tr>
<tr>
<td>Guides to Good Practice</td>
<td>24</td>
</tr>
<tr>
<td>Additional Projects</td>
<td>24</td>
</tr>
<tr>
<td>Publications and Conference Papers by UKDA and HDS Staff</td>
<td>25</td>
</tr>
<tr>
<td>Diary of the Year 2001 - 2002</td>
<td>28</td>
</tr>
<tr>
<td>Feedback</td>
<td>30</td>
</tr>
<tr>
<td>Acronyms</td>
<td>31</td>
</tr>
<tr>
<td>Archive Advisory Committee</td>
<td>32</td>
</tr>
</tbody>
</table>

## Supported by:

- JISC
- ESRC
- University of Essex
Introduction

A year full of achievements, introducing a number of new developments extending the portfolio of services available, the quality of service offered to both depositors and users, and witnessing a marked increase in the volume of activity.

This has been a long year. This statement is true in more than one sense. First, the period covered in the ‘annual’ report spans fifteen rather than the normal twelve months: from 1 October 2001 to 31 December 2002. The reason for this abnormal reporting period results from the fact that the main award from the ESRC, which was due to end in September 2002, was extended by a further three months. The reason for this extension leads me to my next point. Second, during the ‘year’ in question the UKDA was engaged in a prolonged bidding process to continue the range of services it undertakes on behalf of the ESRC and JISC. This process was concluded in the latter part of 2002 and we are pleased to announce a highly successful outcome. Funding has been secured for an initial five years for a new Economic and Social Data Service (ESDS) to be operational from January 2003. ESDS will be managed by the UKDA and run as a joint venture between the UKDA and UK Longitudinal Studies Centre (ULSC) at Essex, and both MIMAS and the Cathie Marsh Centre for Census and Survey Research (CCSR) at the University of Manchester.

A key feature of ESDS will be the development and maintenance of a more integrated approach to data archiving and dissemination, establishing a strategic vision that produces greater interoperability and synergy between the varied ESRC and JISC portfolio of social science data and information resources. In line with this policy, early goals for the new service will be to create a ‘one-stop-shop’ authenticated registration and data access system, together with a social science data ‘universal portal’. Both will aim to provide seamless and easier access to a range of disparate resources.

Within ESDS, the UKDA will continue to operate its central role in the acquisition, preservation, dissemination and promotion of social science data. This will be further augmented by MIMAS providing a new specialist service giving online access to a large collection of international macro time series data. Equally, the CCSR will work to provide data enhancements to a number of heavily-used government surveys, as well as organising a number of specialist training events and user support activities relating to these data.

The UKDA will continue to host a specialist unit for the support of qualitative data (Qualidata) that will place an emphasis on the creation of enhanced, online qualitative data resources within ESDS. The UKDA will also work with the ULSC to both develop and provide a number of data enhancements that will facilitate the use of existing longitudinal data collections. The overall management and co-ordination of ESDS will take place through the UKDA. The new ESDS represents a major investment by the ESRC and JISC, and provides the opportunity to build a world-class resource for the social sciences.

Against this background of bidding and preparation for the new service, the UKDA managed to launch a number of exciting new initiatives and extend both the quality and quantity of services provided for its user community. Fuller details on these activities are provided in the pages that follow. However, it is appropriate to highlight a selection of these achievements in outline form:

- integration of the Qualidata unit into the UKDA;
- launching of a new instant Download dissemination service;
- an increase of 65 per cent in the number of datasets disseminated;
- an increase of 199 per cent in the number of orders serviced;
- an increase of 75 per cent in the number of users serviced;
- an increase of 21 per cent in the number of datasets processed;
- launching of the new Census Registration Service;
- publication of Preserving & Sharing Statistical Materials;
- preparation of a new Digital Preservation Policy.
It is hard, if not impossible, to single out any one of these accomplishments from the string of achievements. However, particularly rewarding is the dramatic rise in the number of distinct users serviced. In each of the four years prior to the period covered by this report, the number of users serviced by the UKDA was relatively static, at around the 1,000 mark. This context makes the 75 per cent rise witnessed in this period even the more remarkable. Moreover, it would seem to indicate that the new Download and Nesstar dissemination services recently introduced by the UKDA have had a significant impact in widening the community of data users within the social sciences. This is a very welcome development.

The period covered by this report also saw the publication of the latest Research Assessment Exercise (RAE) results. This, as most in the UK academic sector will know all too well, is an extensive consultation exercise and peer review process by which each department (or more strictly each Unit of Assessment) in each university in the UK is classified in terms of the quality of its research output. The census period for the latest RAE covered research outputs for the five-year period 1996 to 2000. Linking the ratings of every university department within the UK for the five key social science units of assessment (Business and Management, Economics, Politics, Geography, Sociology) with usage statistics for the UKDA over the same census period produces some interesting findings, as illustrated in Figures 1 and 2, below. Overall, 72 per cent of research-active social scientists in these five key subject groups were recorded in the top three RAE categories (4, 5 and 5*). However, over the RAE census period 90 per cent of the data distributed by the UKDA was to those in departments assessed as being either 4, 5 or 5*. Moreover, a third of all data distributed went to departments judged to be of the highest research quality (5*) even though such departments accounted for only 14 per cent of all research-active social scientists. Thus the clear majority of UKDA users over the period covered by the RAE were from departments judged to have undertaken research of excellence (5 and 5* departments accounting for 62 per cent of use) with those in the highest rated departments ordering an average 20 per cent more data than would be expected given the size of such departments. This exercise is no doubt an imperfect tool, but it does go some way in indicating that the services provided by the UKDA have helped support research of the highest standard.

Kevin Schürer, AcSS
Director, UK Data Archive

---

**Figure 1. Proportions of research-active staff and data supplied by RAE category**

<table>
<thead>
<tr>
<th>RAE category</th>
<th>% staff</th>
<th>% data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3b</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>3a</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>5^*</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 2. Proportions of data ordered under or over that expected by RAE category**

<table>
<thead>
<tr>
<th>RAE category</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3b</td>
<td>3b</td>
</tr>
<tr>
<td>3a</td>
<td>3a</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>5^*</td>
<td>5^*</td>
</tr>
</tbody>
</table>
Services to Depositors

The UK Data Archive aims to position itself as a leader in the field of the management and preservation of data resources by building on the experience and expertise gained from 30 years work, and the drive and enthusiasm to remain at the cutting edge of developments.

The UKDA implemented a number of improvements in its services to depositors over the reporting period. Internal data management processes were integrated and standardised across the core UKDA, Qualidata and HDS specialist units. Data validation, preparation, and preservation processes were improved with the expansion of process guides and the introduction of additional automated scripts for some of the more repetitive tasks.

Development and integration of internal tracking databases across all Archive activity was a further area of development this reporting period. Throughout, the UKDA has striven to become more standardised in its practice and transparent in its procedures. This has enhanced not only the direct services provided to data creators and data users, but has also raised the UKDA’s profile and reputation in the realm of data management and preservation. Consultancies with the Medical Research Council (MRC) and visitors from institutions in France, Netherlands, Czech Republic, Australia, Hong Kong, and America are testimony to this reputation.

Table 1. Datasets acquisition by subject

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>1998-99</th>
<th>1999-00</th>
<th>2000-01</th>
<th>2001-02</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social issues</td>
<td>40</td>
<td>82</td>
<td>90</td>
<td>74</td>
</tr>
<tr>
<td>Labour</td>
<td>36</td>
<td>47</td>
<td>125</td>
<td>38</td>
</tr>
<tr>
<td>Health</td>
<td>16</td>
<td>54</td>
<td>39</td>
<td>32</td>
</tr>
<tr>
<td>Industry</td>
<td>11</td>
<td>9</td>
<td>16</td>
<td>24</td>
</tr>
<tr>
<td>Politics</td>
<td>12</td>
<td>28</td>
<td>25</td>
<td>24</td>
</tr>
<tr>
<td>Economics</td>
<td>42</td>
<td>38</td>
<td>33</td>
<td>21</td>
</tr>
<tr>
<td>Government</td>
<td>15</td>
<td>12</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>Education</td>
<td>10</td>
<td>15</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>Social structure</td>
<td>13</td>
<td>14</td>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td>Media</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Science/Tech.</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Population</td>
<td>7</td>
<td>9</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Social welfare</td>
<td>7</td>
<td>9</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>Travel</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Housing</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Other</td>
<td>5</td>
<td>8</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

KEY ACHIEVEMENTS 2001-2002 AND PROGRESS TOWARDS STRATEGIC PLAN

ACQUISITION

Identifying and acquiring data and other resources according to a collections policy which furthers the strategic goals and mission of the ESRC.

A central activity over the past year has been to improve the support given to depositors and creators of data. The acquisitions team was reorganised and new in-house databases designed to provide improved support and tracking of acquisitions. New, simpler, easier to navigate Depositor Forms were drafted, and a new rationalised Deposit Licence written for implementation in 2003.

In addition, the process of appraising potential acquisitions has been re-examined, and a new Acquisitions Review Committee established. In the period covered by this report a total of 226 data submissions were reviewed (excluding HDS), of which 172 were accepted and 54 rejected. In addition a further 69 datasets were accessioned into the UKDA collection, either via the HDS or preceding the implementation of the new Review procedure. (Figure 3 and Table 1.)

Figure 3. Datasets acquisition by source

<table>
<thead>
<tr>
<th>1998/99</th>
<th>1999/00</th>
<th>2000/01</th>
<th>2001/02</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Government</td>
<td>80</td>
<td>85</td>
<td>85</td>
</tr>
<tr>
<td>Academic</td>
<td>40</td>
<td>40</td>
<td>35</td>
</tr>
<tr>
<td>Data Archive</td>
<td>20</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Other</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>
VALIDATION & CONVERSION

Validating the data and documentation, and converting them to machine-readable formats appropriate to archival and dissemination needs.

New working practices were consolidated this year and procedural guides expanded. In addition, new tools were developed to standardise and automate data validation and conversion tasks. Following the launch of the new instant Download dissemination service (see Services to Users section), data are now routinely converted to multiple user formats, typically SPSS, STATA and ASCII tab delimited (with additional variable type and label information). Work also continued on preparing the most popular data series for publication via Nesstar for online browsing, tabulation, subsetting and downloading. Forty-four new datasets were published in Nesstar over the reporting period, bringing the total to 58 as of 31 December 2002. Overall the number of datasets processed in the reporting period was 622, a rise of 21 per cent from the figure for 2000-2001 (Figure 4).

Significant work was also undertaken in legacy areas, to improve and migrate old data resources as appropriate for user demand (see Preservation and Collections Management section).

METADATA CREATION

Producing information about the data resources in order to populate finding aids including the catalogue and indexes, the web site and other dissemination tools.

The year saw the initial development of new tools for catalogue metadata creation, which will be launched in 2003. These new tools will improve efficiency of metadata creation, and will ensure interoperability with all the data catalogues of the sister social science data archives in Europe under the new Integrated Data Catalogue (IDC-2, see entry for MADIERA project in the Research and Development section). These tools are completely DDI-based and incorporate improved standardised controlled vocabularies to enhance interoperability.

PRESERVATION & COLLECTIONS MANAGEMENT

Preserving data and documentation to ensure they remain usable over time and reviewing the collection to ensure appropriateness.

The Archive pursued an active preservation agenda during this period led by an assessment of its preservation procedures and policies by an international expert. The outcomes were a ‘good housekeeping seal of approval’, a written preservation policy adopted by all staff, heightened security measures, updated internal procedures, and a focused agenda for the next two years.

Further support of this core activity was achieved by collaboration with various professional organisations that resulted in a preservation service for the Archaeological Data Service, a subject centre of the AHDS; membership of the new Digital Preservation Coalition; and the production of a joint publication with the Royal Statistical Society, Preserving & Sharing Statistical Material. This booklet was launched in March 2002 and has subsequently been both well received and widely distributed.

One further strand of this activity was the organisation of a day conference for regional archivists investigating the ‘Challenge of Electronic Records’, held in Cambridge.
<table>
<thead>
<tr>
<th>Planned Tasks Outlined in Strategic Plan</th>
<th>Progress in 2001-2002</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Validation &amp; Conversion</strong></td>
<td></td>
</tr>
<tr>
<td>■ Implementation of the four level model for processing data, to ensure that effort is placed on those data which are expected to have the most productive usage, whilst preparing others for preservation.</td>
<td>■ Levels of processing are now well established, with incoming and outgoing levels assigned at the time of acquisitions review. This has resulted in resources being maximised where they are most effective.</td>
</tr>
<tr>
<td>■ The continued improvement of staff productivity and the integration of data and metadata services.</td>
<td>■ Staff productivity continues to improve, and staff cross-skilling is ongoing. Major gains have been made in closer working among Depositor Services and User Support and Dissemination staff.</td>
</tr>
<tr>
<td>■ Development of more packaging and value addition of the data in ways to make it more suitable for use by secondary analysts.</td>
<td>■ Nesstar publication continues apace for the most popular data series, and all new datasets are now mounted on the instant Download service, access conditions permitting.</td>
</tr>
<tr>
<td><strong>Metadata Conversion</strong></td>
<td></td>
</tr>
<tr>
<td>■ Implement the DDI DTD metadata standard, including the four level model for datasets.</td>
<td>■ Achieved.</td>
</tr>
<tr>
<td>■ Rationalise and improve the thesaurus.</td>
<td>■ ELSST multilingual social science thesaurus continues to be improved, and will be implemented as part of the planned IDC-2 CESSDA catalogue under the EC-funded MADIERA project in 2003-4.</td>
</tr>
<tr>
<td><strong>Preservation &amp; Collections Management</strong></td>
<td></td>
</tr>
<tr>
<td>■ Review the collection and remove or improve datasets as required.</td>
<td>■ Collections management tracking and procedures were formalised and improved with the establishment of a ‘legacy hotline’ in January 2002. As a direct result, 319 datasets were improved and a further 683 new user guides for legacy datasets were created.</td>
</tr>
<tr>
<td>■ Continue to review security issues and implement improvements as necessary.</td>
<td>■ Key improvements to physical storage conditions for all media and proactive software and hardware security upgrades to assure long-term preservation of Archive materials throughout their life cycle were achieved.</td>
</tr>
</tbody>
</table>
Services to Users

The UK Data Archive is committed to assisting users to locate, acquire, access, and analyse data resources as quickly and easily as possible; either from within its own holdings and services, or through collaborative efforts and links with sister data and support organisations internationally.

This year saw a significant enhancement in terms of allowing users to gain instant access to many of the UKDA’s most popular data holdings. In December 2001, the UKDA’s instant web-based Download service was launched, thereby complementing the fully functional Nesstar service for online browsing, tabulation, subsetting and download of data. A number of the most heavily-used datasets, including many of the major series, are now available for immediate download. Downloadable data are available in three standard formats: ASCII tab delimited, SPSS and STATA.

Census users also saw noticeable gains in how they access UK census data. In August 2002 the Census Registration Service (CRS), based at the UKDA, launched its one-stop shop for those in UK higher or further education. Users can now seamlessly gain access to census data resources from 1971 to 2001 using Athens Single Sign On, instead of, as previously, having to navigate through multiple registration hoops. Further details of the CRS project are given in the Research and Development section of this report.

The key driver behind these developments is the wish to provide simple and harmonised methods for users to access and use a range of disparate data resources. These concerns feature high on the agenda of the new ESDS, that aims to provide a fully joined-up data provision and support service for the research, teaching and learning communities.

KEY ACHIEVEMENTS 2001-2002 AND PROGRESS TOWARDS STRATEGIC PLAN

DISSEMINATION

Providing user access to data and documentation, via the web or fixed media, taking account of their technical requirements and reformatting data when necessary.

Figure 5. Delivery times

<table>
<thead>
<tr>
<th>Year</th>
<th>Instant</th>
<th>&lt; 1 week</th>
<th>1-2 weeks</th>
<th>2-4 weeks</th>
<th>&gt; 1 month</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997/98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998/99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999/00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000/01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001/02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Usage this past year rose dramatically as a direct result of the introduction of the instant Download service. UKDA delivery times are therefore now heavily skewed towards instant delivery. Ninety-eight per cent of orders were delivered within the stated target of 4 weeks compared to 93 per cent last year (Figure 5). The number of datasets delivered, increased dramatically from 8,723 last year to 14,295, an increase of 64 per cent (Figure 6). Likewise the number of orders serviced shot up from a level of 1,933 to 5,780, a remarkable increase of some 200 per cent. Particularly rewarding was the substantial increase recorded in the number of distinct users serviced. Over the last 4 years this figure had plateaued at around the 1,000 mark. However, in the current recording period this increased to 1,736 users, a rise of some 75 per cent. This last statistic would indicate that the introduction of the new download service has significantly widened the community that the UKDA serves.

Table 2. Twenty most popular data series

<table>
<thead>
<tr>
<th>RANK</th>
<th>TITLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Quarterly Labour Force Surveys</td>
</tr>
<tr>
<td>2</td>
<td>General Household Survey</td>
</tr>
<tr>
<td>3</td>
<td>Family Expenditure Survey</td>
</tr>
<tr>
<td>4</td>
<td>British Household Panel Survey</td>
</tr>
<tr>
<td>5</td>
<td>British Social Attitudes Survey</td>
</tr>
<tr>
<td>6</td>
<td>Eurobarometer Survey Series</td>
</tr>
<tr>
<td>7</td>
<td>1970 British Cohort Study (BCS70)</td>
</tr>
<tr>
<td>8</td>
<td>Health Survey for England</td>
</tr>
<tr>
<td>9</td>
<td>Workplace Employee Relations Survey</td>
</tr>
<tr>
<td>10</td>
<td>1981 Census</td>
</tr>
<tr>
<td>11</td>
<td>Family Resources Survey</td>
</tr>
<tr>
<td>12</td>
<td>British Election Studies</td>
</tr>
<tr>
<td>13</td>
<td>Youth Cohort Study of England and Wales</td>
</tr>
<tr>
<td>14</td>
<td>National Child Development Study and BCS70</td>
</tr>
<tr>
<td>15</td>
<td>National Child Development Study</td>
</tr>
<tr>
<td>16</td>
<td>Northern Ireland Family Expenditure Survey</td>
</tr>
<tr>
<td>17</td>
<td>Road Accident Data</td>
</tr>
<tr>
<td>18</td>
<td>Continuous Household Survey</td>
</tr>
<tr>
<td>19</td>
<td>British Crime Survey</td>
</tr>
<tr>
<td>20</td>
<td>ONS Omnibus Surveys</td>
</tr>
</tbody>
</table>

Note: The NS Databank, although a very popular series, is not reported in this table for two reasons: first is the measurement of user accesses which is not recorded in a comparable way to other major series; and second, in 2002 ONS launched their own significantly improved online service for these time series and the UKDA now directs users to this site.

RESOURCE DISCOVERY

Assisting people in locating relevant data, and developing and implementing tools and technology to improve resource location across organisational and national boundaries.

The UKDA continued to build on its strength in building high quality resource discovery tools with the implementation of a new and improved search catalogue that includes expanded capabilities for variable, time period and thesaurus-aided searching. At the same, a more global vision was pursued to bridge international boundaries. UKDA information specialists worked with members of sister data archives to establish a European implementation of the Data Documentation Initiative (DDI) metadata standard with associated controlled vocabularies for key elements.
Continual enhancements were made to the UKDA web pages in response to user suggestions. Streamlined navigation features for ‘finding data’ and ‘online data’ and geographic-focused web pages all served to optimise resource discovery. On the technical side, W3C standards were implemented to the same end.

**SUPPORT**

*Supporting users of data, acting as a gateway or filter between users and depositors, supporting data creators in the deposit process.*

During the period 2001-2002 a total of 3,512 queries were recorded into the query tracking database. Post-order queries doubled to 652 from last year’s figure of 332, largely due to the significant increase in usage via the instant Download service. Pre-order queries make up the majority of User Services enquiries, accounting for a total of 2,860. Pre-order queries arise primarily from users seeking advice on the nature and content of datasets.

Of post-order queries, three-quarters were resolved within a day (compared to two-thirds the previous year), while 91 per cent were resolved within one week. For pre-order queries, 93 per cent were resolved in a day and 98 per cent within a week.

**PROMOTION**

*Promoting the wider use of Archive services and providing an information and link point to other relevant resources.*

The newsletter of the UKDA, UK DAtabytes, is published quarterly with a circulation of some 4,500.

The User Services section continued to undertake a programme of outreach and promotional work to both inform, engage and support users. Six user-oriented workshops were held over the reporting period, which enabled both novice and experienced users to gain first-hand knowledge about accessing and using key data resources. The UKDA was also active in contributing to a wide range of social science events, in which new online resource discovery and dissemination services were promoted.

The User Services staff continued to actively feed into training and teaching initiatives. In addition to contributing to the final report for the JISC-funded project on the Use of Numeric Data in Teaching and Learning, the UKDA won an award for a new JISC project on re-purposing data in teaching and learning, under the Exchange for Learning (X4L) Programme. Work on the award will start in the coming year.

<table>
<thead>
<tr>
<th>Planned Tasks Outlined in Strategic Plan</th>
<th>Progress in 2001-2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dissemination</td>
<td></td>
</tr>
<tr>
<td>■ Implementing the Nesstar data and metadata browsing system.</td>
<td>Achieved.</td>
</tr>
<tr>
<td>■ Implementing the Nesstar data and metadata subsetting and delivery system.</td>
<td>Achieved.</td>
</tr>
<tr>
<td>■ Implementing new online registration and ordering system.</td>
<td>Achieved.</td>
</tr>
<tr>
<td>■ Implementing new simplified charging policy.</td>
<td>Achieved. In order to move in line with the ESDS service, charges were waived for all users in the HE/FE sectors who were undertaking non-for-profit research.</td>
</tr>
</tbody>
</table>
- Revamping redistribution agreements and expanding network of redistribution partners.
- In the past year online centralised download services provided by the UKDA has meant that redistribution agreements per se are less of an issue. Consequently the focus has been on enhancing local networks of support for data users rather than dissemination points.

- Developing and implementing interoperable user authentication.
- Single Sign On using Athens authentication has been implemented for the Census Registration Service. One-stop registration will be further extended to meet the needs of the new ESDS.

- Improving and functionally merging internal administrative databases.
- All of the UKDA's internal databases have been undergoing a major review and redevelopment to promote cross-divisional interoperability. The catalogue input programmes have been successfully migrated from Ingres to SQL server.

### Resource Discovery

- Developing the next generation of resource discovery and online ordering tools, including a web-based BIRON replacement with some Nesstar functionality and a 'shopping basket' approach to ordering online.
- Achieved.

- Expand contextual metadata.
- A suite of geographic pages launched for Scotland, Northern Ireland and Wales to provide alternative access to key data collections.

- Implement a multilingual thesaurus for search and retrieval.
- LIMBER project completed successfully and MADIERA project launched. The latter will extend the ELSST thesaurus to include Finnish, Greek, Norwegian and Danish.

### Support & Promotion

- Developing web FAQ database and expanding support web pages.
- New web pages have been designed on ‘using data’ focusing on supporting users. These provide links to training resources across the UK.

- Producing online promotional slide shows for download by reps, lecturers and research officers.
- Achieved. In addition, all workshop materials are made available for download. More training slide shows in pipeline.

- Participating in social science events.
- UKDA contributes to social science events on a regular basis, including being present at forums and providing packs for key conferences and workshops.
Research and Development

The development of new services and the exploitation of new technologies remain a major division of UK Data Archive activity – in this fast-moving world the Archive is well positioned to lead or contribute to national and international developments.

As in previous years, research and development activities have helped to keep the UKDA at the forefront of the data services world. Over this reporting period the UKDA has both completed and started a number of projects. These are summarised below. Such projects, often undertaken in national or international partnerships, are fundamental to the UKDA’s strategic aims in that they allow technical and procedural developments to be undertaken alongside the core activities of data acquisition, preservation, dissemination and user support. Once fully developed, these can then be mapped across and rolled-out as part of the core UKDA service.

In addition to the specific projects outlined below, it is equally important to note the core development work done by the internal programming, preservation and systems, and information development teams. These teams continuously carry out vital work on the internal ‘backend’ systems that enable the smooth flow-through of datasets from data creator to data user. Over the past year much has been accomplished: further refinements to the internal tracking databases; new functionality for the main UKDA catalogue; a new data Download service. During this period the UKDA also commissioned a thorough review of its preservation policy by an international specialist in digital preservation. This has not only confirmed the robustness and integrity of the systems previously in place, but has enabled a revised preservation policy to be written and implemented, taking account of changes in technology and current views on best practice in terms of records management.

PROJECTS ENGAGED IN DURING 2001-2002

CENSUS REGISTRATION SERVICE (CRS)

Under the 2001 Census Programme, jointly funded by the ESRC and JISC, the UKDA was awarded a contract to host a new central registration service for the census products emanating from the 2001 UK census. The central aim of this project, which runs for five years from August 2001 and is co-ordinated by Lucy Bell, is to provide an integrated, seamless, user-friendly one-stop registration service for access to all the varied census resources from the 1971, 1981, 1991 and 2001 decennial censuses. Much of the first year of the project was spent in consultations with potential users and stakeholders, as well as basic systems developments, Athens implementation and beta testing. The system went live according to schedule in August 2002 and since then has been managing the registrations for the four Census Data Support Units funded under the 2001 Census Programme: the Census Dissemination Unit at MIMAS; the Census Geography Data Unit (UKBORDERS) at EDINA; the Census Interaction Data Service at the Universities of Leeds and St Andrews; and the Samples of Anonymised Records from the Census Microdata Unit at the CCSR. From launching the new service until the end of December 2002, a total of 4,911 users had registered on the new system from a total of 240 HE and FE institutions from across the UK. For further details of the CRS see www.census-registration.ac.uk.
**CHCC PROJECT**

This collaborative project between MIMAS, the Leeds School of Geography, the CCSR, the HDS, the LTSN Subject Centre for History, Classics and Archaeology and the UKDA is being funded by JISC between August 2000 and September 2003. The central aim of the project is to increase use of the Collection of Historical and Contemporary Census Data in learning and teaching.

The HDS and the LTSN Subject Centre for History, Classics and Archaeology are making the Historical Censuses Collection more accessible by disseminating it as an online integrated package of data, primary and secondary texts and images, student guides, exercises and teacher guides. The end product will be suitable for the study and teaching of many aspects of nineteenth-century British studies.

An online data delivery system containing data from the 1881 Census for England and Wales, the Channel Islands and the Isle of Man, the National Sample from the 1851 Census of Great Britain and the Great Britain Historical Database will go live in Spring 2003. Owing to the possibility of users being confused by the fact that there is only aggregate-level census data for Ireland, it has been decided to add data from the Database of Irish Historical Statistics at a later stage when users are accustomed to the content of the system.

**MRC DATA ARCHIVING & ACCESS PROJECT**

In recent years the MRC has funded the construction of a large number of population-based studies. This includes several longitudinal studies and a significant number of cross-sectional studies and clinical trials. But to date the MRC has not had a formal policy with regard to the archiving and provision for secondary analysis of these data. To that end, the MRC Data Archiving and Access Project was established in 2001 to gather information, consult widely, and at the end of the project, to make recommendations to Council concerning data archiving and access policy.

Phase One of the project conducted a broad but general survey and convened a working group of interested experts. Phase two tightened the focus by commissioning a series of site visits to collect in-depth information on the conduct of population-based data creation and management; and equally broadened the focus by convening a ‘Horizons’ workshop which attempted to locate the current inquiry in the broader context of developments in e-science generally.

The UKDA played a key role in both phases, first consulting informally with the MRC as they were formulating their broad survey; then sitting on the Data Archiving Working Group which considered the results of the first survey; and finally being hired as consultants to help conduct the Phase II site visits. The results of the study were presented in a substantial consultants’ report at the ‘Horizons’ workshop in October 2002.

As a result of this study and the UKDA’s recommendations, the MRC has formally adopted the principle that data resources created with MRC funding should be made available for secondary use. It has further issued a call for pilot projects in data archiving and access, which will help to further shape MRC policy in this area, and investigate different frameworks for data management and data provision.
FASTER

The FASTER project, funded by the European Commission, ended during 2002 following a successful final review and commendations from the reviewers. They were particularly impressed with the establishment of Nesstar Ltd. to exploit the IP from the project, and recommended the project for an IST ‘best-managed project’ award.

COSMOS

COSMOS is an EPROM cluster designed to bring together a number of organisations working on EC projects to develop statistical software. The UKDA was invited to join, based on the successful work of the FASTER project. The general aims of the cluster are to exchange information between the project partners and to demonstrate interoperability between the different projects. The basis for this work is the architecture designed for FASTER. A second requirement for interoperability is the need for a common model for official statistics. Each of the projects has its own model and a significant amount of work has already been done in mapping these and identifying those elements required in a common model. Again the work of FASTER features highly since the DDI, implemented in FASTER, is at the heart of the COSMOS model. Further details are available at www.data-archive.ac.uk/home/cosmos.asp.

METANET

MetaNet is an EC network under the 5th Framework Programme. It brings together experts from both Europe and North America, in different areas of metadata modelling including classification models and models for micro and macro data. The UKDA was invited to join as an associate member, contributing expertise in metadata on behalf of data archives and data disseminators. Further details are available at www.data-archive.ac.uk/home/metanet.asp.

AMRADS

AMRADS is an EC accompanying measure designed to bring together discrete work in a number of areas relating to the collection and dissemination of official statistics. The measure incorporates a number of strands, for example statistical disclosure control and metadata. The UKDA was invited to contribute expertise in metadata, specifically in the context of user needs. Further details are available at www.data-archive.ac.uk/home/amrads.asp.

GEO-SPATIAL PROJECTS

In 2001 the Geo-data Browser feasibility studies recommended that work in this area should continue. In response, the JISC awarded further funding to the HDS at the UKDA, and EDINA, at the University of Edinburgh.

The aim of the new award has been to build a geo-portal demonstrator to aid the discovery of spatially referenced data. A mock-up has now been created and subjected to early testing within the UKDA. In January 2003, the first external demonstrations will take place as part of a workshop to be held at Essex. Further evaluation will then be undertaken at external organisations.

The GeoXwalk project complements the portal project. It will create a scalable, online gazetteer service, designed to support geographic searching in the JISC Information Environment. Although at this stage the service will be limited, it will be capable of extension to a full service and will act as proof of concept. In line with the recommendations of the feasibility study, this project focuses on near-contemporary data. Further details are available at www.data-archive.ac.uk/home/geoxwalk.asp and www.data-archive.ac.uk/home/gogeo.asp.
MADIERA

The UKDA is part of a consortium of CESSDA partners which has recently been awarded funding from the EC to develop an effective infrastructure for the European social science community. This will be achieved by integrating data with other tools, resources and products of the research process. The end of the project will see the integration of the UKDA into a fully operational web-based infrastructure, populated with a variety of data and resources from a selection of providers. MADIERA began in December 2002. Further details are available at www.data-archive.ac.uk/home/madiera.asp.

METADATER

EC funding has also been awarded as part of the MetaDater project which will begin early in 2003. This project will focus on the development of a data model and metadata management tools for survey research. In addition, it will consider the technical aspects of harmonising survey data from diverse sources and will contribute to work on standards for the description and management of large-scale comparative surveys over time and space. Further details are available at www.data-archive.ac.uk/home/metadater.asp.

ARCHIVES HUB

This JISC-funded project ran for ten months from March 2002 under the leadership of Louise Corti. Its purpose was to contribute collection-level content to the JISC-supported Archives Hub service, based at MIMAS. The Archives Hub provides a single point of access to descriptions of archives held in around 60 UK universities and colleges. The collection descriptions are being compiled from the holdings of the four major national archival resources based at Essex: the UKDA, Qualidata, HDS, and the Albert Sloman Library Special Collections. Further details are available at www.archiveshub.ac.uk.

COLLABORATION WITH THE RSS

Over the course of the reporting period the UKDA worked closely with the Royal Statistical Society to produce a publication to promote the benefits of preserving and sharing statistical material. This booklet *Preserving & Sharing Statistical Material*, is now available and has been circulated widely to interested organisations worldwide.
<table>
<thead>
<tr>
<th>Planned Tasks Outlined in Strategic Plan</th>
<th>Progress in 2001-2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>■ Develop improved metadata models to handle more data types.</td>
<td>■ Extensive work undertaken in this area by the FASTER, COSMOS and MetaNet projects, and will be extended as a result of the new MADIERA and MetaDater projects.</td>
</tr>
<tr>
<td>■ Develop multilingual thesaurus services.</td>
<td>■ The LIMBER project, which focused on this goal, was completed in December 2001. The tools from this project will be extended and implemented through the new MADIERA project.</td>
</tr>
<tr>
<td>■ Merge BIRON and Nesstar developments for data discovery.</td>
<td>■ This process is already underway and will develop further with the MADIERA project. The BIRON catalogue has already been replaced.</td>
</tr>
<tr>
<td>■ Provide more flexible clients for data browsing.</td>
<td>■ Developments of the Nesstar suite of software and extensions to its functionality are ongoing. A new version of Nesstar will be implemented in the coming year.</td>
</tr>
<tr>
<td>■ Develop improved access control services.</td>
<td>■ The CRS launched a new Athens-compliant registration system. This will be broadened to cover all UKDA services over the coming year.</td>
</tr>
</tbody>
</table>
Resource Management

The Resource Services section is committed to providing administrative support to meet the objectives of the UK Data Archive.

Over the reporting period the UKDA has, through the University of Essex, employed a total number of 78 staff. Of these, just under a half of the staffing component were employed part-time (35 staff), with a number also being employed on fixed, short-term contracts, particularly in the case of those working on projects (see Research and Development section). Given both the numbers of staff employed and the degree of change and turnover, the function that the Resource Services team plays is vital.

An Activity-based Costings exercise detailing how staff time is spent and allocated within the UKDA started in the previous reporting year and was continued into this. The extensive monitoring process provided a comprehensive set of activity costs. These have enabled, and will continue to enable, the UKDA to plan, manage and allocate the resources available with greater efficiency and effectiveness. In particular, the figures proved especially valuable for the bidding process and costings exercise for the new ESDS.

KEY ACHIEVEMENTS 2001-2002 AND PROGRESS TOWARDS STRATEGIC PLAN

QUALITY MANAGEMENT

Developing existing staff and recruiting new skills into the Archive by extending the reach of the quality system to include a Career Appraisal System to manage the development of existing staff.

Work continues on implementing the Archive’s Staff Development and Review scheme. A set of procedures for the formal review of all staff has been fully revised and will be implemented in the coming year. This process will identify skill gaps within the UKDA’s staff component and address these with training as appropriate.

Over the reporting period the UKDA has continued its registration with Reach Online, an online training service specialising in online training, skill assessment and development. A number of staff participated in the scheme. In addition, the UKDA, jointly with the Staff Development Office of the University supported four members of staff undertaking external courses, mainly modules with the Open University.

Work has also continued on generating procedures to cover the various aspects of the Archive’s work. These continue to be updated to ensure that the Archive is meeting all its requirements and that it falls into line with the University’s rules and regulations.

PROJECT MANAGEMENT

All Archive projects will be managed according to PRINCE methodology ensuring that they meet the strategic objectives of the Archive, are driven by and judged according to user needs and utilise the best available technology.

The Resource Services section continues to assist UKDA staff in all aspects of project management. Externally-funded projects in which the Archive is currently involved are detailed under the Research and Development section of this report. The section also provides valuable assistance to UKDA staff on the preparation and submission of new project and funding proposals, as well as project reporting.

All Archive projects continue to be managed according to PRINCE methodology.

PLANNING & RESOURCE MANAGEMENT

Integrating all of the Archive information systems to provide better decision-making information.

Progress continues on the integration of the Archive’s information systems. Further functionality has been added to the query tracking database and its interoperability with the order tracking system (STORM) has been improved. The first version of a new acquisitions and data processing database (CALM) has been implemented and is undergoing further modification and refinement.
FINANCIAL SUMMARY

The UKDA’s budget continues to become increasingly diffuse in terms of the sources of funding that it receives. Over the reporting period the Archive received some £1.85 million across twenty separate awards. The accompanying Venn diagram illustrates that for the period 2001-02 the core ESRC award accounted for 48.9 per cent of total funding. This compares with a figure of 51 per cent for the previous reporting year. However, it should be noted that the UKDA also received funding from the ESRC for both the Qualidata and Census Registration Service, jointly 8.7 per cent of the funding total. EC-funded projects dropped in importance, in financial terms, from 16 per cent of total funding last year to just 3.9 per cent in the current period. This situation is likely to change in the next reporting period due to two new EC projects going on stream.

<table>
<thead>
<tr>
<th>Planned Tasks Outlined in Strategic Plan</th>
<th>Progress in 2001-2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>■ Designing and introducing more effective management information systems to more easily provide information on the status of Archive activities and costs.</td>
<td>■ Achieved. Activity-based Costs were successfully collected.</td>
</tr>
<tr>
<td>■ Setting up a modular management information system with all activities of the Archive linked in a series of co-ordinated databases.</td>
<td>■ A new acquisitions and data processing database (CALM) implemented and further functionality added to the query tracking database and STORRM.</td>
</tr>
<tr>
<td>■ Enhancing the existing support of senior Archive staff in contract negotiation and preparation.</td>
<td>■ Achieved. Resource Services staff provide required support.</td>
</tr>
<tr>
<td>■ Identifying the skills gaps for filling by existing staff or through external recruitment to introduce new skills.</td>
<td>■ New procedures for staff development and review written. Online training courses in place.</td>
</tr>
<tr>
<td>■ Develop existing staff to provide consultancy and teaching services.</td>
<td>■ Over the past three years new staff have been recruited with a skills profile that helps facilitate this objective. In the past year the UKDA has undertaken a major consultancy for the MRC (see Research and Development section) and senior staff have had input into national consultations in Australia, Canada, France and Switzerland on the establishment of new data archiving facilities.</td>
</tr>
</tbody>
</table>
In October 2001, the Qualitative Data Service (Qualidata) formally merged with the UKDA to become a specialist unit based in the UKDA. Prior to that, since 1994 when it was first established, Qualidata existed as a separate unit within the Department of Sociology at the University of Essex.

The drivers behind merging the data services were multifold: the desire to create a one-stop social science data shop built around a single hub; the need to strengthen alliances to meet a tendering process ensuing from the ESRC’s strategic review of their data archiving and dissemination services; the wish to streamline and simplify the data deposit process for ESRC depositors; and a growing need to reduce the demarcation between qualitative and quantitative data.

Due to preparation work in advance of the formal merger date, phase one of the integration process was complete by October 2001, with most of the strategic and operational procedures for data acquisition, processing, metadata creation and dissemination in place. Staff were both physically and structurally fully integrated within the UKDA infrastructure by early October.

The period up to December 2002 saw further efforts to harmonise working practices. First, a programme of cross-divisional training was initiated to broaden the data processing skills of UKDA staff to cover a wider range of data types, including mixed methods datasets. Second, the Qualidata web site and online catalogue were transferred to the UKDA servers. Finally, Qualidata rolled out a programme of work to create freely-available online user guides for all its major collections.

NEW WEB SITE

A new look Qualidata web site was launched in February 2002. The new site has been redesigned to parallel the basic navigational features of the UKDA web site, with sections on news, finding, using, creating and depositing data.

An extended section on data resources offers advice on ways of using qualitative data, pointers to key articles on methodological and analytic issues surrounding re-use, and links to other key data resources. New to the site is a section on international data archives, which charts the progress of other data centres around the world in the process of establishing Qualidata–like initiatives.

ACQUISITION & DEPOSITING SERVICE

Enabling researchers to fulfil their contractual obligations to ESRC. Accessioning, processing, archiving and documenting significant datasets generated by ESRC researchers.

Qualidata continues to run an efficient enquiry and advice service for ESRC grant applicants, data creators, ESRC or otherwise, and data depositors. Over the reporting period Qualidata received 121 enquiries relating to ESRC awards, up from last year’s figure of 105. Around three-quarters of enquiries related to grant applicants seeking advice on a variety of topics: costing and preparation of data for archiving; ethical issues of data confidentiality and...
informed consent; and the availability of data in particular areas, as prompted by the ESRC award application form. The remaining queries were from award holders seeking clarification on the depositing process. As part of the Qualidata service, ESRC award holders and Programme Directors are offered advice at an early stage to facilitate the eventual archiving of the qualitative data arising from their research project.

As part of the Qualidata service, ESRC award holders and Programme Directors are offered advice at an early stage to facilitate the eventual archiving of the qualitative data arising from their research project.

As part of the merger with the UKDA, Qualidata also participated in the newly-established acquisitions review process, whereby all data offered are assessed for suitability from both long-term usability and technical perspectives.

This period, 65 datasets were evaluated, representing a slight reduction compared to the 77 offered last year (Figure 7). This is most likely due to the fact that the ‘annual chase’ of award holders traditionally carried out in the spring of each year did not take place. Of these 65, which were mostly ESRC-funded, 26 were rejected, primarily on the grounds of being too small or too narrow in focus, while 39 were accepted. The graph below shows the breakdown of data evaluated by discipline.

**Figure 7. Number of studies evaluated by subject**

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Number of Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sociology</td>
<td>15</td>
</tr>
<tr>
<td>Socio-legal studies</td>
<td>10</td>
</tr>
<tr>
<td>Social policy</td>
<td>5</td>
</tr>
<tr>
<td>Social anthropology</td>
<td>5</td>
</tr>
<tr>
<td>Psychology</td>
<td>5</td>
</tr>
<tr>
<td>Political studies</td>
<td>5</td>
</tr>
<tr>
<td>Management studies</td>
<td>5</td>
</tr>
<tr>
<td>Linguistics</td>
<td>5</td>
</tr>
<tr>
<td>Geography</td>
<td>5</td>
</tr>
<tr>
<td>Education</td>
<td>5</td>
</tr>
<tr>
<td>Economics</td>
<td>5</td>
</tr>
<tr>
<td>Area studies</td>
<td>5</td>
</tr>
</tbody>
</table>

This period, ten digital datasets were processed and deposited at the UKDA. It is worth highlighting two of these new deposits to indicate the wealth of information they contain.

**Conflicts and Violence in Prison, 1998-2000.** This study addresses the interpretation and control of violent incidents within the contemporary penal system. The project has a mixed methods foundation comprising structured questionnaire surveys of a sample who were participants in, or witnesses to, violent incidents in prison. Further qualitative research was carried out using ‘incident analyses’ of violent, or potentially violent, incidents, whereby demographic and biographical details of prisoners involved, extracts and summaries from interviews held with prisoners, witnesses and prison staff concerning the incident, were recorded.

**Cultural Industries and the City, 1980-1999.** This study explores the artistic and cultural development of Manchester during the 1990s. Using semi-structured interviews with cultural entrepreneurs working within the key sectors of music, fashion, multimedia and design, the research builds up a picture of contemporary working practices in an unexplored area of British life, and addresses the contribution such production and activity makes to shaping both regional and national identity.

Qualidata facilitated the deposit of a further four paper-based qualitative collections in traditional repositories.

**SERVICES TO USERS**

**Disseminating information about archived qualitative data and monitoring the re-use of data.**

In the past, Qualidata has had the difficult task of attempting to measure usage of data held in a number of archival repositories (e.g. special collections). Records that are kept are rarely comparable, if they are collected at all. As such, usage figures in the past were likely to under-represent the true patterns. This year usage has been measured only of studies deposited at the UKDA and at the National Social Policy and Social Change Archive (NSPSCA), a paper-based archive housed in the University of Essex Library. The NSPSCA is a rich and unique resource comprising the records and data from many of the earlier ‘classic’ social science studies that employed (but not exclusively) qualitative methods.
Qualidata User Services staff logged 495 queries about using data in the reporting period. Queries covered questions about whether data exist on a particular topic or theme, availability of data from specific projects, data content, and how to access and cite data.

The period saw 133 users of 9 data collections archived at Essex. Looking at re-use in more depth, qualitative data are used in four ways: research training; social history research; comparative research; re-use data and research instruments for research design and piloting.

Nevertheless, the fact remains that a quarter of the potential users enquiring about data, do not go on to locate a dataset suitable for their needs. In other words, the supply does not, by any means, meet the demand, since the majority of qualitative datasets created in the UK are simply not archived nor made available by investigators for secondary use.

Qualidata has found that usage figures tend to reflect the relative promotional efforts invested in these datasets. An example is Paul Thompson’s *Family Life and Work* interview data, which has been re-used by a very large number of researchers over the past 20 years, in contrast to the very similar and potentially equally valuable interviews from his *Families and Social Mobility* project, which has gained almost no explicit publicity. Another very well-publicised high-use example is the Mass Observation material at Sussex, which typically attracts some 800 or so users per annum.

All new qualitative datasets that do not have special conditions of access attached to them, have now been mounted in the UKDA’s instant Download service. Easy access to online qualitative data and user documentation represents a significant improvement for users who, in the past, may have had to travel to a far-off repository to use a set of transcripts. Instead, users now have the advantage of being able to ‘click and go’, and within a few minutes can have data downloaded to their desktop to import into their favourite CAQDAS package.

A next obvious step, in terms of improving access to data, is to enable users to search and browse the content of data online, in a manner akin to exploring survey data via the Nesstar software. To this end, in December 2001 the UKDA initiated a new project aimed at developing a framework and system for providing online access to qualitative data. Phase One of the *Edwardians Online* project was completed in May 2002, culminating in the launch of a pilot digital multimedia resource that integrates a wealth of existing primary and secondary materials based on a large-scale oral history project. The first formal evaluation exercise was conducted in September 2002, resulting in both positive feedback and helpful suggestions for Phase Two of the project. Further details of the *Edwardians Online* project can be found at www.qualidata.essex.ac.uk/edwardians.
PROMOTION

Disseminating information about depositing, sharing and using qualitative data.

Running promotional and training events, and creating training resources is one of the principal ways of encouraging re-use of data. Indeed, these tasks are essential for the continued success of Qualidata’s mission, and can also be viewed as an integral part of operating a datasets policy. The Diary of the Year section of this report demonstrates that Qualidata staff have both organised and been present at a number of key promotional events. This period has seen Qualidata venture into new territory, with two workshops being organised at the request of local FE colleges. The growing interest in the use of data within pre-University teaching and learning is an area being encouraged and facilitated by the JISC.

Qualidata staff were invited to advise on a number of budding initiatives across Europe, centred on building qualitative data centres, including Switzerland, France and the Netherlands.

Qualidata has a regular feature in the UKDA quarterly newsletter, UK DAtabytes, which has a circulation of some 4,500.

---

<table>
<thead>
<tr>
<th>Planned Tasks Outlined in Strategic Plan</th>
<th>Progress in 2001-2002</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acquisition &amp; Depositing Service</strong></td>
<td></td>
</tr>
<tr>
<td>■ Advising ESRC current and prospective award holders on data creation issues of project design, costing and data management</td>
<td>■ All enquiries and requests for further information are answered as a matter of priority.</td>
</tr>
<tr>
<td>■ Advising award holders on the preparation of data for electronic archiving.</td>
<td>■ Deposit web pages improved to guide depositors through the data preparation process.</td>
</tr>
<tr>
<td>■ Proactive and co-ordinating support for ESRC Programmes relating to qualitative research.</td>
<td>■ Acquisition staff continue to reach out to Programme award holders by visits to Programme meetings and sending out packs.</td>
</tr>
<tr>
<td>■ Maintaining an information and dissemination service and up-to-date documentation.</td>
<td>■ Revised ‘create and deposit’ web pages and guidelines.</td>
</tr>
<tr>
<td>■ Contributing to professional bodies’ Codes of Ethics for research in relation to archiving.</td>
<td>■ Contribution to working parties of the SRA and BSA on revising their codes of ethics.</td>
</tr>
<tr>
<td>■ Acquiring datasets and preparing licence agreements for deposit in the UKDA.</td>
<td>■ Achieved and ongoing. The deposit of key qualitative collections that reside in paper format in traditional repositories, is facilitated by Qualidata.</td>
</tr>
<tr>
<td>■ Processing and preparing metadata for no more than 15 key ESRC electronic datasets of a qualitative nature.</td>
<td>■ Achieved.</td>
</tr>
<tr>
<td>■ Playing a role in data documentation initiatives and their applicability for qualitative data within the data archive and traditional archival communities.</td>
<td>■ Edwardians Online project started work on building a framework and standards for a qualitative data metadata and data exchange format.</td>
</tr>
</tbody>
</table>
### Services to Users

- Providing an up-to-date information service about available sets of data.
- Evaluating delivery of qualitative electronic datasets via the UK Data Archive mechanisms.
- Supporting user enquiries about availability and use of data.
- Qualidata online catalogue maintained, web pages on NSPSCA collection, and browsing facility by discipline.
- Achieved. All new datasets processed are deposited at the UKDA and, where suitable, mounted on the UKDA’s instant Download service. Online delivery systems are being developed.
- User support staff answer and track all user queries.

### Promotion

- Maintaining dedicated JISCmail group.
- Input into workshops jointly with UKDA.
- Contributing to social science organisation’s regular newsletters.
- Membership now stands at 332 members.
- Qualidata staff contribute to workshops and participate in external events on using and depositing qualitative data. See Diary of the Year.
- Qualidata staff are both invited and offer to publish in relevant social science outlets, such as newsletters and edited collections. See Publications section.
- Users are contacted where personal details are available. ESRC Regard database is regularly surveyed for publications.
The History Data Service (HDS), recently renamed the AHDS Subject Centre for History, is a specialist unit located within the UKDA. As one of the service providers of the AHDS it is funded jointly by the JISC and the AHRB. It was originally founded in 1993 and became part of the AHDS in 1996. It reports separately through this body, but a brief report of highlights is presented here in order to present a rounded picture of the UKDA’s work as a whole.

It is also appropriate to note that Cressida Chappell, who had been both Acting Head and Head of the HDS since 1999, left the HDS in November 2002 to take up a new position at the University of Oxford. A new Head of Service, Matthew Woollard, formerly of the History Department at the University of Essex took over from 1 January 2003.

HIGHLIGHTS OF THE YEAR
- Prototype download service providing access to thirty-six data collections.
- HDS web site enhanced to validate to HTML 4.01 Transitional and meet the requirements for Priority 1 of the W3C Web Content Accessibility Guidelines 1.0.
- Funding to provide data storage for the NOF-funded Citizenship Past Project.
- Funding for 12 months for the Go-Geo! Geo-Data Portal Project (Phase 2).
- Funding for 12 months for the GeoXwalk Gazetteer Project (Phase 2).

RESOURCE DELIVERY
A prototype download service that allows users to select and download data and documentation when they require it in suitable easy-to-use formats was launched in November 2002. Thirty-six data collections are currently available, but in the longer term most of the data collections held by the HDS will be made available to users via this new system.

TRAINING WORKSHOPS
The HDS held the Designing Flexible Digital Representations of Historical Source Materials workshop at Essex in December 2001. The workshop discussed the methodology of creating digital representations of historical source materials. During the workshop participants were guided through some of the key issues involved in the conversion process. The workshop articulated and demonstrated those standards and elements of good practice that are relevant to a wide range of data creation situations, from student projects through to large-scale research projects. The workshop presented an overview of the issues involved in the design of digital representations of historical sources. Reasons for, and benefits of, creating digital representations were considered, as were typical scenarios and problems. Particular emphasis was placed on the importance of a systematic approach and thorough documentation. Workshop participants were provided with pointers to paper-based and online literature that discusses the issues addressed in this workshop in more detail.

WEB SITE
The HDS web site has been enhanced to validate to HTML 4.01 Transitional and meet the requirements for Priority 1 of the W3C Web Content Accessibility Guidelines 1.0. New additions include a short introductory guide to planning historical digitisation projects which was web-published in April 2002. The guide highlights important issues that anyone planning to create a digital resource for research learning or teaching in history, or an associated discipline should address. Topics covered include: project management; project timetables; monitoring progress; acquiring
GUIDES TO GOOD PRACTICE

A Place in History: A Guide to Using GIS in Historical Research was web-published on 12 September 2002 to coincide with the AHDS GIS in the Arts and Humanities workshop in York. It was written by Ian Gregory and peer reviewed by an editorial board consisting of Roger Kain, Donald Morse, Alasdair Crockett, Alastair Pearson and Graham Jones. A more traditional book version will be published in April 2003 with Oxbow Books. It is intended for historians who want to use GIS and 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.

This Guide to Good Practice 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.

ADDITIONAL PROJECTS

The HDS has managed and participated in many projects throughout the reporting period. Of these, the CHCC, Archives Hub and Geo-spatial projects are summarised under the Research and Development section of this report due to their collaborative nature with UKDA.

Citizenship Past Project

The HDS is participating in this large-scale digitisation project funded under the New Opportunities Fund’s Digitisation of Learning Materials Programme. The project, which is led by the British Official Publications Collaborative Reader Information Service (BOPCRIS) based at Southampton University, also involves the West Yorkshire Archives Services, the Children’s Society, MIMAS, based at Manchester University, and the Centre for Data Digitisation and Analysis based at Belfast University. The HDS is responsible for storing the 500GBs of scanned documents that will be produced by the project.
Publications and Conference Papers by UKDA and HDS Staff

Five issues of the UKDA’s newsletter UK DAreadyes produced and published. ISSN 1474-9149.


October 01
Qualidata unit merges with UKDA.
Kick-off meeting for EC-sponsored COSMOS project hosted by UKDA.

November 01
Joint workshop held with Centre for Longitudinal Studies on Using the National Child Development Study & 1970 British Cohort Study, London.

December 01
Workshop on Using the Scottish Household Survey, held with Scottish Executive, Edinburgh.
UKDA’s new data Download system launched.

January 02
UKDA staff attend Data Documentation Initiative Meeting, Washington DC
Elizabeth Lyon, Director of UKOLN visits UKDA.
Launch of EEDA Observatory Data Module using Nesstar software.

February 02
Director invited as ‘external expert’ to partake in consultation exercise on creation of a new Canadian Data Archive.
Mattanja de Boer visits UKDA from the European Foundation, Dublin.
Deborah Mitchell and Trevor Breusch, Director and Deputy Director of the Australian Social Science Data Archive visit on fact-finding tour.
Myron Gutmann, new Director of ICPSR visits UKDA.
Melanie Wright attends Digital Preservation Coalition launch, House of Commons.

March 02
Final review of EC-funded FASTER project, nominated for IST ‘best managed project’ prize.
Publication of Preserving & Sharing Statistical Material by Royal Statistical Society and UKDA.
Francoise Cribier and Elise Feller from CNRS, Paris, visit Qualidata.

April 02
UKDA staff present poster session at Geo-data forum, Manchester University.
Qualidata staff present papers at a workshop on archiving qualitative data, Neuchatel, Switzerland.

May 02
UKDA organise day meeting, held in Cambridge, on the Challenge of Electronic Records for regional archivists.
Melanie Wright attends DLM Forum 2002 in Barcelona.
Lorna Balkan presents paper on EC-sponsored LIMBER project at Third International Conference on Language Resources and Evaluation, Canary Islands.
**June 02**

UKDA and HDS staff present a total of eight papers at the IASSIST 2002 Conference, *Accelerating Access: Enhancing Collaboration and Dissemination*, Connecticut.

UKDA host RDN *Train the Trainer* workshop.

BBC film Qualidata unit for Open University programme, *Trouble with Love*.

**July 02**

David Lewis, EU consultant on data dissemination visits UKDA.

UKDA staff attend JISC workshop for Service Providers, *Collection and Service Description*.

**August 02**

New Census Registration Service launched.

**September 02**

Qualidata staff present papers at DRH 2002, Edinburgh University Library.

Karl Wilding, NCVO, visits UKDA.

Peter van den Besselaar (new Director of Netherlands data archive) and Cor van der Meer visit UKDA.

Party of archivists from the Czech National Archives visit UKDA to learn about digital preservation.

Director invited to present paper at CEIES Conference, *Innovative solutions in providing access to microdata*, Lisbon.

Louise Corti delivers SPICE training workshop, Edinburgh.

**October 02**

UKDA and Qualidata staff attend CESSDA Expert Seminar in Romania.

UKDA staff present papers at MRC Horizons Workshop on Population Data Archiving and Access, London.

Diane Geraci attends ERPANET workshop, Urbino, Italy.

Working group from the newly established French Quetelet centre visit UKDA.

**November 02**

Louise Corti gives presentations to JIBS workshop, London.

Party of local FE students attend workshop on UKDA.

**December 02**

Diane Geraci and Louise Corti attend IJSSC/UNESCO conference and IFDO meeting, Vienna.
Feedback...

‘I have been on sick leave for more than half a year. When I came back ...I was ... told that not much has really happened. But when I looked at Edwardians Online I realized that ... a lot has happened. I really look up to you. Your project is inspirational.’

‘Thanks a lot for your help. I look forward to receiving [the data] and please pass on my thanks to the rest of the staff at the Archive for the excellent service you provide.’

‘Thank you very much for the speedy delivery of the CD. Please pass my thanks on to the team for the excellent service you provide.’

‘Thank you for this dataset - downloaded successfully. A very efficient and helpful service all round.’

‘It’s wonderful to get all this information – I’m no expert in this area, and you’ve increased my knowledge enormously.’

‘I would like to say thank you for sorting out my e-mail error, I am impressed at the speed my query was dealt with.’

‘May I say what a helpful and detailed reply that was - and fast! I speedily obtained click-and-use licence and am now attempting to locate the data I need.’

‘Thank you very much for responding so quick and taking action. If only other content providers would be as helpful.’

‘Just a comment about the new online registration. Very efficient. Very clear instructions. Very easy. Much appreciated. Thank you.’

‘Thank you very much for your super-fast reply. I am highly impressed.’

‘Throughout the year, you and your colleagues have offered great help and please let me take this opportunity to say a very big THANK YOU.’

‘It is great to get such a quick response.’

‘Thank you for yet another great experience with your services!’

‘Many thanks for your help ...[I] appreciate your impressive work’

‘You’ve gone beyond the call of duty. I greatly appreciate it.’

‘Thank you very much for answering my query so completely.’

‘Please forward my sincere thanks to you and your colleagues for last week’s session on Qualidata. Our students were immensely grateful.’

‘The [UKDA] home page introduces a professional, well-designed house style, which is maintained consistently throughout the site.’ [from JISC Web Audit conducted by the Digital Media Access Group]

‘In controlling the language for keywords and for subject headings we decided to use the UNESCO thesaurus. However, it doesn’t have much detail for our purposes... [then] I discovered the HASSET thesaurus, which has been extremely useful in amending and placing terms, with a high level of specific examples.’ [from the Content Development Manager, Sense of Place London project, Westminster City Council]

‘Congratulations on another excellent Annual Report. The document is superb: it’s a well-written and informative record of your achievements during the year. The number of goals you reached as outlined in the report is truly impressive. When we revamp our Annual Report next year, I hope to emulate yours – both in style and in substance.’ [from ICPSR]
Acronyms

AHDS  Arts and Humanities Data Service
AMRADS  Accompanying Measure to Research and Development in Official Statistics
AHRB  Arts and Humanities Research Board
ASCII  American Standard Code for Information Interchange
BIRON  Bibliographic Information Retrieval Online
BOPCRIS  British Official Publications Collaborative Reader Information Service
BSA  British Sociological Association
CAQDAS  Computer Assisted Qualitative Data Analysis Software
CESSDA  Council of European Social Science Data Archives
COSMOS  Cluster Of Systems of Metadata for Official Statistics
CNRS  Centre National de la Recherche Scientifique
CSR  Centre for Social Research
DDI  Data Documentation Initiative
DLM  Données Lisibles par Machine (Machine-readable data)
DTD  Document Type Definition
EC  European Commission
EDINA  Edinburgh Data and Information Access
ELSST  European Language Social Science Thesaurus
EPROS  European Programme of Research in Official Statistics
ESDS  Economic and Social Data Service
ESRC  Economic and Social Research Council
FAQ  Frequently Asked Questions
FASTER  Flexible Access to Statistics, Tables and Electronic Resources
FE  Further Education
GIS  Geographic Information System
HASSET  Humanities and Social Science Electronic Thesaurus
HDS  History Data Service
HE  Higher Education
HTML  HyperText Mark-up Language
IDC  Integrated Data Catalogue
ICPSR  Inter-university Consortium for Political and Social Research
IFDO  International Federation of Data Organisations
IP  Intellectual Property
IST  Information Society Technologies
JISC  Joint Information Systems Committee
LIMBER  Language Independent Metadata Browsing of European Resources
LTSN  Learning and Teaching Support Network
MADIERA  Multilingual Access to Data Infrastructures of the European Research Area
MIMAS  Manchester Information & Associated Services
NCVO  National Council for Voluntary Organisations
NESSTAR  Networked Social Science Tools And Resources
NS  National Statistics
NSPSCA  National Social Policy and Social Change Archive
RSS  Royal Statistical Society
SPSS  Statistical Product and Service Solutions
SQL  Structured Query Language
SRA  Social Research Association
UKDA  UK Data Archive
UKOLN  UK Office for Library Networking
UNESCO  United Nations Educational, Scientific, and Cultural Organisation
W3C  World Wide Web Consortium
With the establishment of the new ESDS in January 2003 the UKDA’s Advisory Committee will cease to exist. As part of the ESDS governance plan, the Advisory Committee will be replaced by a new committee with a changed constituency and a refocused strategic outlook. The UKDA would like to extend its thanks to all those who have served on its Advisory Committee over the past years. They have given extremely valuable advice and assistance. Special thanks are due to Professor John Fox who in his position as Chair over the past two years gave much-valued support and guidance during the difficult period of the ESDS bidding process.