CAST: Centre for Applications Software and Technology

by A. Stacey[1]
CAST
Edinburg, UK

CAST/EUL Data Library Services

The Data Library at the University of Edinburgh offers the research community on-line access to facilities for the analysis of census, time–series, sample survey, and related data. These data are held on the Edinburgh Regional Computing Centre's network of ICL 2900 and VAX mainframe computers, in an environment that is well suited to (remote) multi–access interactive computing, and is rich in software for statistical analysis and graphic display. The 'official statistics' data holdings of the Data Library include:

> A very full range of (small) area statistics relating to the 1971 and 1981 Population Censuses for Scotland, England and Wales;

> A series of Parish summary and grid

square data from the Annual Agricultural Census, for England and Wales;

The CSO Macroeconomic Databank:

The Scottish Input/Output Tables for 1979;

The OS Gazetteer;

The Postcode Directory for Scotland;

A full range of digitised boundary files.

Information about the holdings of the Data Library are held in an on–line view system, DATALIB. DATALIB and the holdings themselves are accessible on–line by users from the university communities of Edinburgh, Glasgow and Strathclyde, and by a range of users across the PSS/JANET computer networks, or across the telephone network (using an acoustic coupler/modem). In providing an on–line data library service, the Data Library relies upon the expertise, facilities and staff of the Edinburgh Regional Computing Centre (ERCC), the Edinburgh University Library (EUL) and the Centre for Applications Software and Technology (CAST). It also draws heavily upon the past experience of what was known as the Program Library Unit (PLU).

Access to the BUSH mainframe on the ERCC Network (EDNET) is as follows:

PSS : 2342 313 54354 (and then CALL BUSH)

JANET: 0000 0700 1004 04 (and then 1500 0003 or CALL BUSH)

Datel: (UK) 031–667–1071

[In reply to the prompt 'User' enter

DATALIB]

The Data Library and PLU

The Data Library formally came into being in the mid-1970's in order to provide computing facilities and access to small area statistics from the 1971 Population Census (Scotland). The machine–readable data from this Census were purchased for Edinburgh University's academic research and teaching community by the University Library. The data were managed and were made accessible on the ERCC's network of mainframe computers through programs written by the staff of (what was then called) the Program Library Unit (PLU). PLU was a specialist unit within Edinburgh University with a national role for the conversion and maintenance of statistical packages for ICL computers in universities. It had also gained considerable experience in writing census–access software and had familiarity with computer–aided mapping from work on the annual Agricultural Census (England and Wales) which it had carried out in association with the Department of Geography. (The mapping programs, CAMAP and GIMMS are in widespread use today.) Four programs were written for accessing the 1971 Population Census, together with DATAPAC, a user-friendly interface designed with the novice computer user ('whose expertise lies outwith [sic] computing') in mind.

The PLU played a significant role, in the U.K., in the dissemination and promotion of secondary analysis of (small) area statistics from the 1981 Population Censuses. First, the application software that was commissioned by LAMSAC (Local Authorities Management Services Advisory Council) for the retrieval and manipulation of machine–readable tabular output from the Census was designed and written (under subcontract to Durham University) by staff of the PLU. The SASPAC project, as it was called, won the British Computer Society's Social Benefit Award, and is widely used by local government officers and academic research staff.

Second, PLU played an active part in the two consortia which, on behalf of the academic community, purchased the 1981 Population Census statistics from the offices of the two Registrars–General. In particular, PLU formed the consortium to purchase the data from the Scottish Census, and arranged for these to be deposited with the ESRC Data Archive for general distribution. Later, as part of the Inter-University Software Committee (IUSC)'s working party on census data, the Data Library at Edinburgh became one of the six regional and national computer centres to act as a census data library for academic research purposes. The General Register Office (Scotland) also granted a Census Agency Agreement to the University of Edinburgh to enable the Data Library to provide services to commercial users, including academics conducting contract research and policy analysts in central and local government. PLU has also collaborated in academic research projects. These have included the computerisation of a 2% sample of the enumerators books from the 1851 Population Census, and provision for public access to these data through, for example, the distribution of floppy disks for viewing on the BBC microcomputer.

CAST, EUL and ERCC

The Centre for Applications Software and
Technology (CAST) was established in 1983 as
an outcome of a review of the Program Library
Unit, occasioned by the retirement of its
Director. CAST now has a staff of about 40:
in addition to what are now referred to as the
Data Library Service and the Program Library
Service, CAST also has expertise in database
design and management, numerical algorithms,
statistical computing, graphics, application
software evaluation, conversion and development
(for mainframe and micro computers) and
survey methodology. (CAST continues to
provide some software services under the name
PLU.) The Data Library can therefore call
upon specialists for advice or project work on a
range of relevant activities. In particular, the
provision of access software with attractive user
interfacing (which has for example been written
for the CSO Macroeconomic Databank, the
Agricultural Census and the OS Gazetteer) is
central to the development of the Data Library,
which is why the service is housed at CAST.

The Library of the University of Edinburgh
(EUL) is one of the major university libraries in
the U.K. It has about 200 full-time and
part-time staff, and a large reference collection
including official statistics and maps. It is
currently undertaking the on-line cataloguing of
its holdings, and makes extensive use of on-line
bibliographic search facilities (including
DIALOG, ERIC, etc.). It has begun to grasp
the nettle of cataloguing machine-readable data
files, and many of the computer terminals
installed in the library for use by its 'users' are
terminals also connected to the local computing
network. Both the official statistics and the
map collections provide necessary and
complementary services to the Data Library.
For example, researchers wishing to use the
small area statistics from the Census may
consult the maps in the map reference section

in order to discover where enumeration district
boundaries lie, for example, and consult the
published tables from the Census in the
statistical reference section prior to conducting
an analysis of the machine-readable small area
statistics, and perhaps producing a high quality
schematic map of their own.

The Edinburgh Regional Computing Centre
(ERCC) was founded in 1966, and has
maintained on its network of mainframe
computers an operating system called the
Edinburgh Multi-Access System (EMAS),
specifically designed to provide an interactive
computing environment for a scattered
populaton of users. This was considerably in
advance of the 'star' network systems which
were later developed at national and regional
sites. The ERCC therefore has considerable
experience in both interactive and batch
computing, the provision of on-line help
information and experience in communications
between different mainframes. The EMAS
operating system can also claim to be a
particularly 'friendly' operating system, when
compared to the alternatives available, especially
for the first time user from another computing
site.

General Institutional and Technological Environment

The University of Edinburgh, with over 1500
teaching and research staff and over 10,000
students, has a widely dispersed campus. It is
partly for this reason that Edinburgh has a
telecommunications network that is arguably one
of the most advanced among the universities in
the U.K. The Edinburgh Multi-Access System
(EMAS) has provided a powerful but 'friendly'
interactive and batch computing environment
since 1972. EMAS resides, as an operating
system, on the ICL 2988 and dual ICL 2976

mainframe computers, the latter being connected to two ICL Distributed Array Processors (DAP), and has just been introduced on a newly installed Amdahl mainframe. The EDNET network is a multi-mode packet switching network linking EMAS with VMS and UNIX operating systems on a range of mini-computers, and providing access facilities for the staff and students of Strathclyde and Glasgow universities. The network also offers access to a central (disk) filestore, printing and graphical facilities, an electronic mail service, and a (regional) gateway to the British Telecom Package Switching System (PSS) and the U.K. academic network (JANET). This is shown in the figure below. ERCC is now experimenting with integrated speech and data facilities on its network, evaluating the merits of early access to the full ISDN (Integrated Service Digital Networking) being piloted by BT in London, Manchester and Birmingham.

Current Staff

The staff of the Data Library Service are drawn from the University Library (EUL) as well as from CAST. Staff with major responsibilities, and who regularly spend part of their time in this area, include a manager, a senior computing officer, an administrative/computing assistant, a reference librarian and three computing officers, the latter with specialist responsibilities for programming, computer-assisted cartography and spatial analysis. These staff have skills in database management, statistics, survey analysis and design, cartography and have experience working in the social sciences, government, and the physical sciences.

Work-in-progress

The Data Library at Edinburgh is currently undergoing a reorganisation in order to provide an expanded range of services. This is partly in preparation for the launch of a Scottish Data Centre, but is also partly in response to a wish to foster data library development more generally within the UK. The cataloguing of machine-readable data files has been identified as a priority area, along with the provision of an on-line union catalogue of data collections, indicating the existence and location of data. Some preliminary discussions have been held with Sue Dodd and with the staff of the ESRC Data Archive who have made considerable progress in the compilation of MARC-compatible study descriptions of the Archive's data collections. We also recognise that the development of data libraries depends crucially upon the existence of a national data clearing house, with secure long-term funding, and upon agreement on the data library/clearing house relationship. We therefore have an interest in promoting an organisation like the Data Archive at Essex.

One of the features of Edinburgh's Data Library is the provision of user-friendly interfaces to access software. We have also become conscious of the value of friendly interfaces to mapping packages, and are in the process of designing EASYMAP which will, for example, generate the commands to create a GIMMS map, using digitised boundary data from a machine-readable library.

Funding for these activities comes from two major sources. First, the University provides finance for a core of staff. Second, CAST is able to generate revenue from various external activities. These include, for example, data processing and access management for the Edinburgh District Council's 'Homeless Survey'; consulting to the Scottish Office, and analysis of
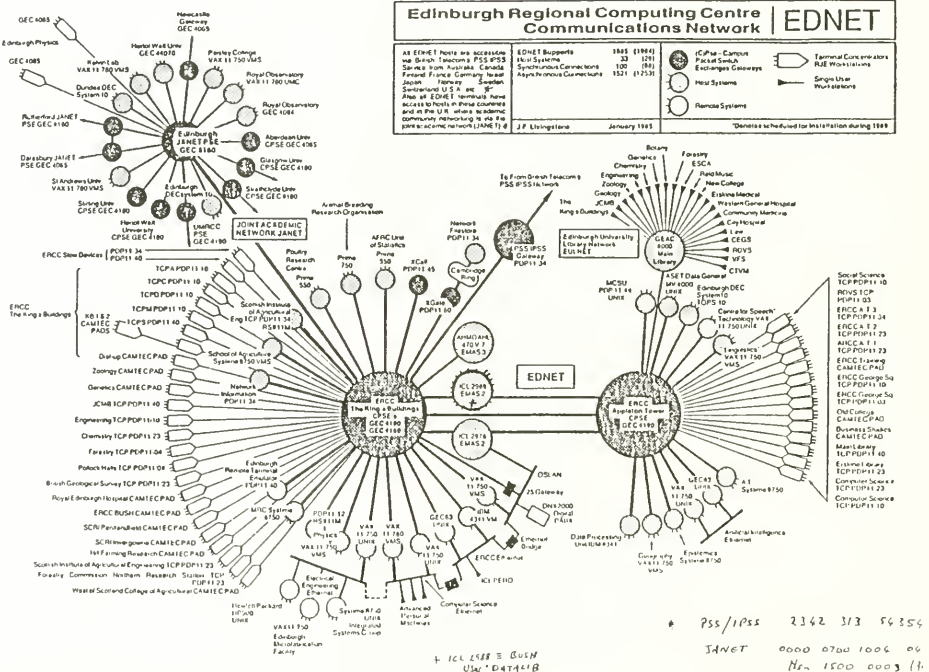
the Agricultural and Populaton Censuses. We are also to seek grant-funding for the development of a more 'public' data library.

Enquiries

Persons who would like to know more about the Data Library at Edinburgh, or who are interested in fostering the development of data libraries in the U.K. are very welcome to

contact Peter Burnhill, Manager, Data Library, CAST, 18 Buccleuch Place, University of Edinburgh, EH8 9LN (tel. 031-6667-1011 exts 6204, 6756). The adventurous might like to try leaving a message on DATALIB, or else sending a MAIL message to DATALIBUK.AC.EDINBURGH across IPSS or JANET.□