

***Geography Data Unit
(UKBORDERS)
Annual Report to the ESRC
Census Advisory Committee
August 2007 – July 2008***

EDINA



EXECUTIVE SUMMARY	2
1 INTRODUCTION	2
2 OVERVIEW OF ACTIVITIES	3
3 PROGRESS	3
3.1 Continued Operation	3
3.2 New Postcode Directory service	3
3.3 Shibboleth infrastructure	4
3.4 User training workshops	4
3.5 Case studies commissioning	4
3.6 Related developments	5
4 INTERNAL COMMUNICATION	5
5 EXTERNAL COMMUNICATION	5
6 FUTURE DEVELOPMENTS	6
ANNEX A STATISTICS	7

Contact Details: James S Reid,
Data Library, Main Library,
George Square,
Edinburgh, EH8 9LJ.

Executive Summary

Bulleted highlights for the years activity are:

- **Continued successful operation of service and support for users;**
- **Development, testing and release of the new postcode data selector service;**
- **Shibboleth infrastructure planning, integration and testing;**
- **Running user training workshops;**
- **Case studies commissioned and released;**
- **Related developments (DIaD project).**

1 Introduction

This report contains a high level review of the activities of the Geography Data Unit (UKBORDERS) during the first period of the new Programme period 2006-2011 for the period August 2007 to July 2008.

Aims

The core aims of the UKBORDERS project are:

- to provide those in UK higher education involved in teaching & research with cost-effective access to a resource of digitised boundaries for thematic mapping and GIS operations; and
- to provide those in UK higher education involved in teaching & research with cost-effective access to geographical lookup directories; and
- to provide training, support and assistance to that community in the exploitation of those resources.

Objectives

The key objectives for the service for the period 2006-2011 are:

- to maintain the data facility as a robust on-line system serving up data in a variety of formats
- to explore and develop new methods and channels to assist users in exploiting the census outputs
- to support users via on-line help, individual assistance and via workshops
- to promote the ease of use of Census materials in research and teaching
- to interoperate with other Census Units to assist in the delivery of census products to the community via the new Census Portal

Start and end dates of the project

The UKBORDERS project commenced on 1 August 2006 and will run until 31 July 2011.

Staffing

The GDU team comprises Dr David Medyckyj-Scott as Unit Director (0.02 FTE). Day-to-day responsibility for the service and ongoing development is the responsibility of James Reid (0.35 FTE). Technical development, GIS analytical support and helpdesk

area. They can further have control over which data items they wish to extract and also the format (CSV, DBF and Shapefile formats are supported). The new service was launched after internal testing on the 3rd July 2008.

3.3 Shibboleth infrastructure

Some of the slippage on the new postcode interfaces has also been a result of the accelerated timescales for Shibboleth introduction and requirements to put in place GDU dependencies in order to test the prototype system at Essex. A series of three way teleconferences between EDINA, MIMAS and UKDA took place over the period December – June in order to realise the practical implementation of Shibboleth integration across all the census resources. Note that GDU plan (in line with other units) to operate from August 2008 a 'dual economy' approach with simultaneous independent support for both Athens and Shibboleth routes. This will maximise user enfranchisement and provide a fallback route for institutions not yet Shibboleth ready although it does imply an additional technical burden for each of the units. Target date for release at time of writing remains 1st August 2008 and internal testing has identified no major technical barriers to completion and smooth roll out as per agreed timescales.

3.4 User training workshops

UKBORDERS was a key feature at three Census Programme workshops run from late November 2007 to early March 2008. GDU and CDU collaborated to organise two workshops at their respective institutions; the third, a Census.ac.uk workshop, was hosted at the University of Cardiff in January. The subject of the joint GDU/CDU workshops centred around combining census aggregate statistics with digital boundary data for mapping and spatial analysis.

The two workshops drew a combined turnout of 40 participants, many of whom attended to gain experience in using Casweb and UKBORDERS data in a Geographic Information System (GIS). The Census.ac.uk geography and migration workshop at Cardiff University workshop attracted 30 people with an interest again in processing UKBORDERS boundary datasets with census data in a GIS to display migration patterns. Feedback from these workshops indicated a strong interest for future workshops including ones with training in advanced boundary and census data GIS applications.

Future workshop collaborations are being planned between GDU and CDU; as well as continued participation with Census.ac.uk at relevant workshops. In addition to revising the future workshop format, GDU and CDU will organise a workshop specifically on the postcode directories, tentatively set to start during autumn term in 2008.

3.5 Case studies commissioning

A series of case studies from the user community have been contributed to UKBORDERS and made available to end users. These case studies provide potential users with an idea of what is possible with the resources available. Further case studies will be released as they are received and reviewed by the support team. The case studies available currently are:

- A study of the Geography of Welsh Speakers in North Wales using GIS
- Epidemiology of Acute Pancreatitis

- Spatial Inequalities: An investigation of spatial disparity in the North-West of England

3.6 Related developments

Under the ESRC Census Innovation awards, GDU were successful in securing funding for a twelve month project, Data Integration and Dissemination (DIaD). The project, in conjunction with CDU, is investigating the potential of using international open standards based techniques to perform data linkage between two of the most heavily used census outputs – the aggregate statistical data and the output geographies. The primary objective of this work is to develop a data dissemination model which demonstrates a more generic capability – that of 'geo-linking'. This provides the ability to separate census statistical data (for example, but other geospatially-linked data are equally capable of utilising this approach) and the boundary (geometry) data to which it relates. Geo-linking allows for distributed, multi-source datasets to be seamlessly linked in a fashion that facilitates data separation for management and administration purposes. In essence, the approach proposed will provide an extensible infrastructure applicable not only to the immediate needs of the Census Programme but also more widely to broader requirements emerging from the e-Social Science programme, especially its e-Infrastructure strand and also the ESRCs national data strategy. Additionally, using the same standards based approach, the project will aim to demonstrate how further value added processing can be invoked by transforming the geo-linked outputs through a series of ancillary web processing services.

4 Internal Communication

As part of the normal operation of the unit various internal (to the census programme) meetings and discussions were held:

- Teleconferences with CRS on CP issues – e.g. metadata, style guide;
- Email exchanges with ONS geography on redistribution of OS derived data sets
- Meetings and email exchanges with OS to clarify data redistribution issues pertaining to derived data
- Teleconferences with CRS on CP issues – e.g. metadata, style guide;
- Email exchanges with ONS geography on redistribution of the ESRC QA enhanced historic postcode resources
- Meetings and email exchanges with OS to clarify data redistribution issues pertaining to derived data and copyright
- Attendance at the initial Data Feeds Network at Leeds at subsequent meeting at GLA in June 2008
- Liaison with MIMAS and UCL for DIaD project

5 External Communication

UKBORDERS, and the Census Access Programme (CAP) more generally, benefit from the promotional activities that EDINA as a JISC National Datacentre engage in. The following summarises the key conferences, workshops and other promotional activities

over the reporting period at which UKBORDERS and the census programme has been promoted.

- Data Citation in "Poverty, wealth and place in Britain, 1968 to 2005" - Daniel Dorling, Jan Rigby, Ben Wheeler, Dimitris Ballas, Bethan Thomas, Eldin Fahmy, David Gordon and Ruth Lupton
- RGS/IBG 2007 Conference, London 28-20 August
- EDINA Newslines, September Edition
- Online Information 2007, 4-6 December, London
- GISRUK 2008, Manchester 2-4 April
- EDINA Newslines, March, 2008 Edition
- JISC Conference, 15th April, ICC, Birmingham
- geoXchange, 23rd April, Dynamic Earth, Edinburgh
- Data Archives meeting, 20th June, Utrecht

6 Future Developments

- The addition of a new Metadata search service. This will exploit the full ISO 19115 / AGMAP [<http://www.gogeo.ac.uk/geoPortal10/Help/Guidelines1.html>] profile metadata records created for the UKBORDERS datasets and will provide users with more informative descriptive details about how datasets were sourced and their technical specifications. Note that this has been delayed due to the need to prioritise the work on Shibboleth integration and the release of the new Postcode Data Selector Service.
- Extensions to the Postcode Data Selector Service to disseminate the historic postcode resources (pre 2000).
- Extensions to the Postcode Data Selector Service to allow fuller user customisation and output flexibility.
- Trialling of interfaces for users to consume ONS Web Services
- Further user commissioned case studies and training activities.
- Development of formal learning and teaching materials.
- Ongoing contributions to the development of the Census Portal.
- General update enhancements to existing services:
 - The implementation of an automated RSS feed service for Census Portal news aggregator and individual user subscription to UKBORDERS news announcements
 - Release of support for KML format in BoundaryDataSelector
 - Additional Projection support in BoundaryDataSelector
 - A redesign and reorganisation of EasyDownload collection.

Annex A – (Summary) STATISTICS

Usage - downloads

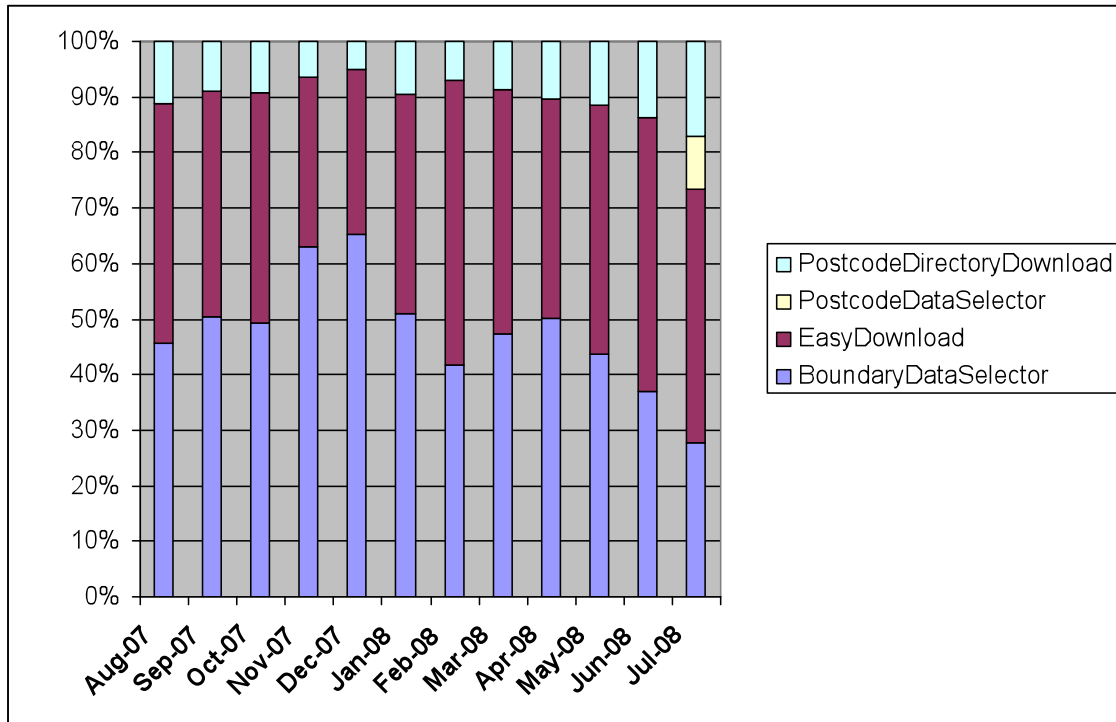
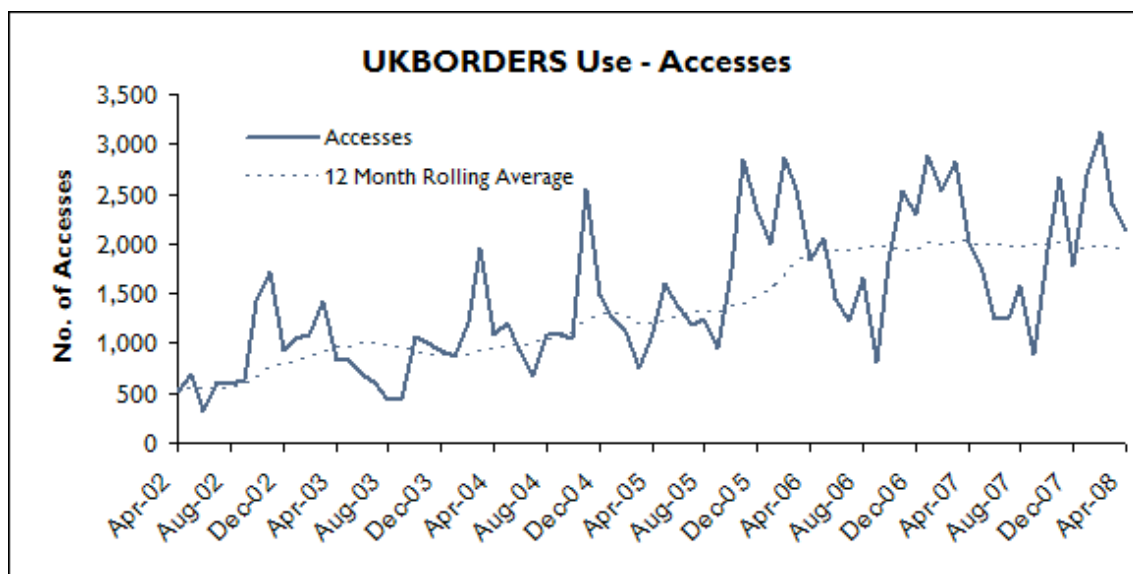


Fig. 1 Download Statistics August 2007 – July 2008.

Number of Accesses - For details on number of accesses to the Census Programme and other DSUs, please see

<http://www.mu.jisc.ac.uk/servicedata/census/trend/index.html>.

The relevant UKBORDERS data is shown below.



	Aug-07	Sep-07	Oct-07	Nov-07	Dec-07	Jan-08	Feb-08	Mar-08	Apr-08	May-08	Jun-08	Jul-08	Total
BoundaryDataSelector	691	422	920	1622	1130	1334	1235	1113	1023	615	395	375	10875
EasyDownload	652	341	771	789	519	1028	1521	1034	804	634	530	614	9237
PostcodeDataSelector	-	-	-	-	-	-	-	-	-	-	-	128	128
PostcodeDirectoryDownload	168	74	171	163	86	252	210	202	213	160	147	229	2075
Grand Total	1511	837	1862	2574	1735	2614	2966	2349	2040	1409	1072	1346	22315

Table 1. BoundaryDataSelector, EasyDownload, PostcodeDataSelector and PostcodeDirectoryDownload numbers of downloads by month, August 2007 – July 2008 [NB PostcodeDataSelector was released publicly on 3/7/08].

Table 2. Top 20 dataset downloads August 2007– July 2008.

Dataset	No.s of downloads
England CAS Wards 2001	2232
England Output Areas 2001	1650
England Districts 2001	1549
England Counties 2001	901
England Super Output Areas (lower layer) 2001	741
England Government Offices for the Regions 2001	499
England Super Output Areas (middle layer) 2001	475
England CAS Postcode Sectors 2001	461
England Urban Areas 2001	372
England Standard Table Wards 2001	362
Northern Ireland Districts 2001	336
England Wards 2001	305
England Wards 1991	310
NSPD February 2008	269
England National Outline 2001	258
England CAS Parishes 2001	257
Scotland Council Areas 2001	241
England National Parks 2001	220
Scotland CAS Parishes 2001	179
NSPD November 2007	174
Total	11791

Helpdesk Calls

UKBORDERS utilises the EDINA helpdesk infrastructure. The majority of calls come through as emails, a few enquiries are by letter, and the remainder are by telephone. Table 3 summarises the type of queries received by the EDINA helpdesk regarding UKBORDERS. The majority of queries are logged and have, as minimum, an initial response within 2hrs. All queries are ultimately resolved and logged as closed by the HelpDesk tracking system.

Short Description	Number of Enquiries
Login or Access Problem	11
Registration	9
Database Contents	65
Interface Use	13
Error Reports	6
Terms	10
Request to View	1
Forward Use	18
Other	6
Total	139

Table 3 Summary of calls received by helpdesk August 2007 – July 2008.