
Let Us Bring You to Your Census: Recent Developments in UK Census Data Provision

This paper describes the development of the one-stop Census Registration Service (CRS), an online system providing quick and simple user registration for access to all the varied resources from the 1971, 1981, 1991 and 2001 UK decennial censuses.

*by Lucy Bell **

conducted every ten years since 1801, with just two anomalies: it was omitted in 1941 due to war-time security issues; and a small-scale, experimental, five-year census was attempted in 1966, using a 10 percent sample of the population.

Introduction

UK higher education institutions, through the UK Data Archive (UKDA) and the Universities of Manchester, Edinburgh and Leeds, have been providing access to UK census data for over a quarter of a century. In that time, the number of data products associated with the 1971, 1981 and 1991 censuses has increased dramatically and now totals more than 50. This number will increase again with the release of the results from the 2001 Census.

Such a wealth of resources presents the academic user with not only the opportunity to use many sophisticated datasets but also a requirement to accept and fulfil the terms and conditions of the various licences covering these products. A new single licence for all 1971-2001 census products is now in development. In line with this new licence and in time for the results of the 2001 Census, the UK Data Archive has been funded by the Economic and Social Research Council (ESRC) and Joint Information Systems Committee (JISC) to develop, implement and maintain the one-stop Census Registration Service (<http://www.census-registration.ac.uk>) to simplify the process of registering for these data products.

This paper reviews the background to the project, explaining the need for such a co-ordinated service in light of the varied actual and potential uses of census data in higher and further education (HE/FE) and the ways in which these applications are currently developing. It also reports the findings of a small-scale questionnaire survey of UK HE/FE staff, undertaken by the paper's author, which sheds light on views of the old and new registration systems. Lastly, the paper describes the new registration service in some detail and highlights some of the lessons learned during its development.

Background

The official census in the United Kingdom has been taken for 200 years, with 2001 being its bicentenary. It has been

The end of the 20th century saw more activity than ever before in disseminating

census materials to academia. In the late 1970s the UK Data Archive became involved, as it started to disseminate the data from 1971. In those days it took nearly ten years to finalise the census data taken at the start of the decade and to publish them. Progress in the following years has speeded this up considerably.

Access for the academic community to these data has also been assisted by a joint ESRC and JISC programme which started with the 1991 Census. This programme funded a series of academic data services, known as Census Data Support Units (DSUs), which have been disseminating the data to the UK higher and, more recently, further education communities in the past decade.

Following the development of these academic services, the data products available to UK HE/FE have grown in number. The 1991 programme encouraged the development of derived datasets and tools which could be used with the census data. There are now over 50 datasets, including resources such as Deprivation Indices derived from the census data, which can be used. These datasets themselves comprise thousands of tables. The geographical breakdown of these data is sophisticated and, for certain data, such as the Small Area Statistics, takes the data to very small output areas. The 1991 ESRC/JISC programme's online and offline access services have also been extended to bring in some of the 1971 and 1981 data. This all gives the UK HE/FE community unrivalled access to a vast collection of census data.

The data essentially break down into the following categories:

- Area statistics
- Boundary data

- Interaction data (origin-destination data)
- Microdata (the Samples of Anonymised Records or SARs)
- Other derived datasets

These data are distributed by the four highly-respected DSUs, funded via the ESRC/JISC Census Programme under its 2001 Census budget:

- Census Dissemination Unit from MIMAS (University of Manchester, <http://census.ac.uk/cdu>) for the area statistics for 1981, 1991 and 2001 and soon 1971
- Census Geography Data Unit (UKBORDERS) from EDINA (University of Edinburgh, <http://edina.ac.uk/ukborders>) for digitised boundary data for 1971, 1981, 1991 and 2001
- Census Interaction Data Service (Universities of Leeds and St Andrews, <http://census.ac.uk/cids>) for origin-destination statistics for 1981, 1991 and 2001
- Census Microdata Unit from the Cathie Marsh Centre for Census and Survey Research (University of Manchester, <http://www.ccsr.ac.uk/sars>) for the 1991 and 2001 Samples of Anonymised Records (SARs)

The combination of the richness of the data resources and the sophisticated dissemination services make the census data valuable for research, teaching and learning; however, one large obstacle has blocked the user's way in past years: registration.

Because the different census products have evolved over time, almost every one has resulted in a new licence being drawn up between the ESRC/JISC and one of the three census offices in the UK. These three administrations organise the censuses for their areas: the Office for National Statistics administers the census for England and Wales; the General Register Office for Scotland administers the Scottish census; and the Northern Ireland Statistics and Research Agency administers the census in Northern Ireland. To complicate matters further, the 1981 and 1991 boundary data for England and Wales and the 1981 boundary data for Scotland are owned and, therefore, licensed by three additional bodies: the Office of the Deputy Prime Minister; a consortium called ED-Line; and the Scottish Executive respectively. Because they have been developed by different organisations, each of these licences is slightly different. Even those licences which relate to the same country can vary in content. Add to this the fact that four distinctly different services distribute the data, all of which have required users to register separately with them, and a complicated system of access becomes apparent.

The original registration procedures to access the census data involved the need to locate, print off, complete and have counter-signed locally, a series of forms, most of which were slightly different in both content and format, in order to be able to use all the census data available. A user approaching the census data in the past may well have had to fill in different forms depending on which census year he or she was interested in, which country was needed and which DSU was being used.

In order to reduce the confusion and the administration associated with having to fulfil each clause of each licence, the ESRC/JISC, the three Census Offices and HMSO have developed a new tripartite agreement which will eliminate all the old licences, bringing the data from each census from 1971 onwards under the same terms. This will create a level playing field, so that all the data may be treated in the same way. The intention is also for this agreement to roll forward to bring in new censuses, as and when they are taken.

The UK Data Archive has developed an online system which co-ordinates all the registrations for access to the census data, after having won the tender from the ESRC/JISC to supply this. It was intended to dispense with the multiple paper forms and create a smooth, user-friendly, Web-based registration system which would knock down many, if not all, of the hurdles in the user's path to the data. The Census Registration Service, established at the UK Data Archive in August 2001, has been developing this system over the course of the past year. The system is now live, meaning that just one online registration now entitles the user to access the data from all the DSUs, for all the years, countries and types of data, using a simple, online, Athens¹-compliant system.

Consultation

One of the first tasks in the CRS's life was to undertake a series of consultations in order to guarantee that the development of the service satisfied all the stakeholders as far as possible. These stakeholders comprised the four DSUs whose registration needs the CRS serves, other experts and a sample of the current system's users.

Three meetings with the DSUs were held during the first year of the service. In order to facilitate confidential and speedy discussions about the progress of the service a closed JISCmail² discussion list was also established in November 2001. The list has been used extensively and has proved to be a vital forum for sharing comments and suggestions.

Consultation with the future users of the service was considered to be just as important as consultation with the DSUs. A series of surveys were prepared and undertaken between November 2001 and March 2002 in order to ascertain the users' views of the current registration

procedures which were then in place. This was done for three reasons:

1. To elicit their views about the then current and forthcoming census registration systems
2. To start to advertise the forthcoming changes
3. To discover whether any of the current users selected would be interested in beta testing the new system

Initially, all MIMAS and UK Data Archive class tutors using census data and UKBORDERS-selected class tutors were contacted with an email questionnaire on Tuesday 4 December 2001 (Appendix A). Class tutors form a special set of census data users. Under the old system, whole classes of students from one university could register in one batch by signing one licence administered by their tutor. This added to the workload of already busy lecturers whose responsibility it became to ensure that their students were signed up to use the data.

The email questionnaire was sent to 48 class tutors; 19 replied. Their responses indicated both familiarity and discontent with the old system:

- 16 said that they either knew how some, most or all of the system worked
- 9 said they found the system either impenetrable or quite difficult to use
- 7 found it 'middling' in terms of ease-of-use
- 10 were dissatisfied
- 2 were extremely dissatisfied

From their written comments, their main concerns were as follows:

- Gathering student signatures is very time consuming
- The registration process involves too much paperwork
- The registration forms tend to be long and complex
- Individual registration for each database is tiresome
- Tutor involvement should be minimised or eliminated

They came up with the following suggestions:

- Speedy online self-registration for students
- A system that is not too technical but easily explained online
- A single form for all census data services

When asked which aspects would discourage them from using – or wanting to use – a system, they came up almost unanimously (n=15) with:

- Tutor involvement in gathering students' names

This was the key. It was reported that it was difficult to organise signatures for an entire class using hard copies of licences because some from the class will always be absent at the time of signature collection. The class tutors commented on their frustration in having to chase the stragglers.

Following on from the class tutors' questionnaire survey, all site representatives from MIMAS, UKBORDERS and the UK Data Archive were emailed the URL of an online questionnaire on 7 February 2002 (see Appendix B for the text of the questionnaire). The CCSR representatives were alerted to the questionnaire through the SARs Newsletter.

In total, 373 representatives were emailed the details of the online questionnaire. Seventy-one replied, yielding a response rate of 19 percent. The results from these respondents mirrored those of the class tutors in that they indicated both an understanding of the system and some dissatisfaction about the way it worked.

- 57.75 percent (n=41) were familiar with at least some of the system, with 32.4 percent (n=23) feeling familiar with most or all of it.
- The respondents did not seem to find the system easy to use. 38.03 percent (n=27) found it 'quite difficult' or 'very difficult' to use, 33.8 percent (n=24) found it 'neither difficult nor easy', with only 9.86 percent (n=7) finding it 'easy'. None of the respondents ticked the 'very easy' box.
- Despite their familiarity with the system, the site representatives were, in the main, dissatisfied with it. 42.25 percent (n=30) ticked the 'extremely dissatisfied' or 'dissatisfied' boxes; 14.08 percent (n=10) were, however, satisfied. No-one ticked the 'extremely satisfied' box.

The reasons supplied for dissatisfaction were:

- The need for counter-signatures on the licence agreements and the reliance on the postal service to deliver them both slow the process down

- The confusing abundance of forms and the difficulty in identifying which forms are needed for which dataset and, indeed, whether all the correct ones have been found, all cause frustration
- The time it takes to register puts off potential users (for students in particular, it was emphasised that the system needs to be speedy)
- The need to remember multiple user names for the services and to register separately for each one is burdensome

The questionnaire included a series of suggested improvements which the respondents were asked to tick should they desire them. Although not all of these additional functions were ticked by each representative, they each received a tick from at least 59% of the questionnaire's respondents. The preferred registration functionality, in order of preference, was as follows:

- A single interface for all services, 95.77 percent (n=68)
- Immediate access to the data after registration, 81.69 percent (n=58)
- Step-by-step online guidance on how to register, 73.24 percent (n=52)
- The functionality to jump from one service to another without logging on again, 66.2 percent (n=47)
- Nothing to sign, 60.56 percent (n=43)
- Helpdesk facility for registration problems, 59.15 percent (n=42)

The results of the class tutors questionnaire were combined with the results of the questionnaire sent to the site representatives and the desires of the respondents incorporated into the system as far as possible.

The functionality of the registration system

After these consultations and further scenario planning, the final specification for the service was decided. This was then programmed, beta-tested and launched. In more detail, the system's features are as follows:

- It is short, simple and straightforward, comprising a wholly online system with no need for paper licences and counter-signatures.
- It is Athens-authenticated, meaning that users will simply have to input their Athens usernames to access the data. This does, of course, mean that users will have to have Athens usernames before approaching the

system; however, this is a straightforward process. In fact, many users will already have Athens identities for use with other resources.

- The CRS verifies users' email addresses using a simple procedure. As soon as the user has completed the online form, a message is sent to the email address they entered. The email contains a URL which they must visit. Once they have successfully completed this, they will have registered to use the census data. There are two reasons for undertaking this email check. First, it confirms the user's email address and, second, it checks that the email goes to the person who completed the registration form.
- The system has a facility for users to update their own details; for instance, if an undergraduate becomes a postgraduate, his or her status may be altered accordingly online.
- An online feedback area exists, where users may submit enquiries to the CRS team.
- There is a facility for inputting the details of publications which have arisen from use of the census data. One of the conditions of the End User Licence is that the user agrees to inform the CRS about these publications. In order to make this easy for the user, the CRS has established a centralised area of the Web site where this may be completed online.
- The system offers registered users the opportunity to sign up for the UK Data Archive as well, removing the need for them to complete additional online forms.
- The site also contains other informative pages, containing news and useful links.

Obstacles and lessons

Few obstacles were encountered during the service's development, but those that did arise were complex. These have now been overcome and the relevant issues addressed. The obstacles essentially fell into two categories: licensing issues and technical issues.

Tracking down all the previous licences, including those for the digitised boundary data which were owned by different organisations than the rest of the census data, proved time-consuming. Additionally, the CRS had hoped to be able to bring other census-related datasets into the registration system in time for its launch, but this proved to be impossible due to time constraints.

The technical obstacles with which the CRS had to grapple related primarily to the different levels of access control that each Data Support Unit required for their data. For example, some units were clear that, should the service

be opened up outside the UK HE/FE community, it would not be possible for their data to be disseminated to these additional users. The CRS was intended to be a one-stop shop but, because of this access issue, it has also been developed with flexibility in mind. This flexibility was achieved through use of the Athens profile system.

The Athens system allows each of the services to be treated differently; each can be set up as a separate 'resource' in the Athens terminology. Each user who is granted access to a resource has an electronic 'profile', to which information may be written and from which information may be read. The census system has been established so that, once a user registers with the CRS, the expiry date of his or her registration is written to the user's profile. When the user logs into one of the DSUs the Athens software reads the profile and assesses whether or not the user is entitled to access the data. This system can be extended to include other information regarding the user's status to allow or disallow access to individual DSUs. For instance, should a user work outside the HE/FE sector, a fact which can be determined from the Athens username, information about access rights can be included in the user's profile, preventing him or her from using some units and allowing them the use of others.

Lessons about how not to re-invent the wheel were learned early on in the project. The use of externally established systems of communication and authentication assisted the CRS enormously. One of these was Athens, while another was the JISCmail closed discussion list. This allowed all parties involved to reflect on changes to the system and to suggest the best solutions. It also resulted in agreement among all four of the DSUs.

Support has also been forthcoming from many quarters. The DSUs have encouraged the service's development and given ideas and advice all the way, as have users in the field. Colleagues in the UK Data Archive have been on hand to discuss various options whenever needed. This has supplied the CRS team with a rich and varied body of experts from whom to take advice.

The launch and the future

The service was launched on 2 August 2002. From this point onwards, all UK HE/FE census data users have registered with the CRS. Previous users have had to re-register with the service for legal reasons. The UK Data Protection Act 1998 means that the CRS cannot simply transfer users' details from one service to another; however, the bonus that exists in re-registration is that the process should be short, simple and straightforward. Most importantly, filling in one online form entitles users in the academic community to use all four of the Census Data Support Unit services and the UK Data Archive.

The CRS is now entering Phase 2 of its life and is gearing

up to enhance and improve the service. The developments planned include a search interface to the database and more links. The CRS team will also be listening carefully to its users' suggestions. If additional changes take place, they could well be a result of comments made by those whom the CRS values very highly – the people using the data.

Footnotes

1 Athens (<http://www.athens.ac.uk>) is a system used throughout UK HE/FE which supplies students and staff with just one username and password to access many different databases. It works with the education – and other – sectors and with database producers. HE/FE institutions must negotiate access for their user communities to different databases, after which the database producers inform Athens which institution's members may use their services. The institutions are then granted access to these databases by Athens. Athens has over 1.8 million user accounts. All UK HE/FE institutions and the Research Councils have already been granted access to the census services, as these are the audiences for whom the services have been funded.

2 JISCmail (<http://www.jiscmail.ac.uk>) is a service based on LISTSERV which allows UK academics and support staff to join, create and manage electronic discussion lists.

*Paper presented at the IASSIST Conference, June 2002, in Storrs, CT, USA. Lucy Bell, UK Data Archive, lajbell@essex.ac.uk.

Appendix A: Email to class tutors, sent to 48 people via email on Tuesday 4th December 2001

Dear Dr ~

The Census Registration Service at the UK Data Archive <<http://www.data-archive.ac.uk/>> is in the process of setting up a one stop shop Census Registration Service, providing a user-friendly and simple registration system for access to all the varied census resources from the 1971, 1981, 1991 and 2001 decennial censuses. This means that, soon, just one registration per person will be required to access all census materials from 1971, no matter where the data are located (MIMAS, UKBORDERS, Centre for Censuses and Survey Research etc.). The service is expected to be ready in the summer of 2002. For more information on the project, please follow the link to the UK Data Archive's projects web pages <<http://www.data-archive.ac.uk/home/CensusRS.asp>>.

My role is to coordinate this process and one of the first things I am trying to do is to canvas opinions on the ideal design and functionality of the new service. Colleagues at MIMAS mentioned your name as a tutor who uses the class registration procedures to access census materials

with your students and suggested that you may have extremely valuable views on how the new system should be developed in this regard. If you had five or ten minutes to answer the questions below and email them back to me, I would be very grateful. Indeed, any comments about the current and/or future registration systems in relation to the needs of class tutors would be gratefully received. If at all possible, it would be helpful to have all responses back by Friday 21st December.

Many thanks.
Lucy Bell
Service Coordinator
Census Registration Service
UK Data Archive, University of Essex
Wivenhoe Park Colchester Essex CO4 3SQ
Tel: 01206 873950
Email: lajbell@essex.ac.uk

**** **Questionnaire: Access to the Census Datasets** ****

1. On a scale of 1 - 5, where 5 is the greatest, how familiar do you feel with the current systems of class registration required to access the 1971-1991 census datasets? PLEASE MARK.

- 1 __ I don't know how the system works at all
- 2 __ I know only a little about how it works
- 3 __ I know some of it quite well
- 4 __ I am fairly sure how most of it works
- 5 __ I know exactly how it works

2. On a scale of 1 - 5, where 5 is the easiest, how easy do you find the current census registration systems to use? PLEASE MARK.

- 1 __ Impenetrable
- 2 __ Quite difficult
- 3 __ Middling
- 4 __ Easy
- 5 __ Very easy

3. On a scale of 1 - 5, where 5 is extremely satisfied, how satisfied are you with the current registration procedures? PLEASE MARK.

- 1 __ Extremely dissatisfied

- 2 __ Dissatisfied
- 3 __ Neither satisfied nor dissatisfied
- 4 __ Satisfied
- 5 __ Extremely satisfied

Please give the reason(s) for your answer to question 3:

- 4. What would make you, as a class tutor, view the new census registration system as a success?
- 5. Which aspects of an online registration system would discourage you, as a class tutor, from using it?
- 6. If there are any anonymised suggestions or complaints you have picked up from students using the current census registration system and which you feel you can share, please describe them below.
- 7. Any other comments:

If you would like to express an interest in the beta testing of the new service in 2002, please just let me know.

Many thanks for your time in answering these questions.

Lucy Bell
Service Coordinator
Census Registration Service
UK Data Archive
University of Essex
Wivenhoe Park
Colchester
Essex
CO4 3SQ
Tel: 01206 873950
Email: lajbell@essex.ac.uk

Legal Disclaimer: Any views expressed by the sender of this message are not necessarily those of the UK Data Archive. This Email and any files transmitted with it are confidential and intended solely for the use of the individual(s) or entity to whom they are addressed.

Appendix B: Online questionnaire for site representatives

Census Registration Service Questionnaire

Many thanks for taking the time to visit this page and complete the questionnaire (below). The results will be stored in a database held at the UK Data Archive. The information is being gathered to help inform the development of the new one-stop Census Registration Service for UK higher and further education.

This new service will provide an integrated, seamless, user-friendly and simple registration system for access to all the varied resources from the 1971, 1981, 1991 and, when they are ready, 2001 decennial censuses. This means that, soon, just one registration per person will be required to access all census materials from 1971, no matter where the data are located (MIMAS, UKBORDERS, Cathie Marsh Centre for Census and Survey Research, CIDS or the UK Data Archive). The service is expected to be ready in the summer of 2002. The answers you give to the questions below will help to determine its final design.

1. How familiar do you feel with the current system of registration required to access the 1971-1991 census datasets?

- I don't know how the system works at all.
- I know only a little about how it works.
- I know some of it quite well.
- I know, reasonably well, how most of it works.
- I know exactly how it works.
- I have never used the system (please go to question 4).

2. How easy do you find the current census registration system to use?

- Very difficult.
- Quite difficult.
- Neither difficult nor easy.
- Easy.
- Very easy.

3. How satisfied are you with the current census registration procedures for UK higher and further education?

- Extremely dissatisfied.
- Dissatisfied.
- Neither satisfied nor dissatisfied.
- Satisfied.
- Extremely satisfied.

Please give a brief summary of the reason(s) for your answer for question 3:

4. If usage statistics for this service for your institution were available, which level of breakdown would you find most useful?

By which category of user? (tick all that apply)

- I would not need the results to be broken down by category of user.
- Educational type (staff, postgraduate, undergraduate etc).
- Department.
- Subject area.

By which time period? (tick one)

- Annual figures.
- Quarterly figures.
- Figures produced more frequently than quarterly.

5. What would you, as a user, like to see in an ideal census registration system? (tick all that apply)

- A single interface for logging on to all services.
- Nothing to sign.
- The functionality to be able to jump from one service to another without logging on again.
- Immediate access to the data after registration.
- Helpdesk facility for registration problems.
- Step-by-step online guidance on how to register.
- Something else (please specify).

6. Additionally, please indicate on a scale of 1-5 where 5 is the most satisfied, your satisfaction with the following aspects of the UK Data Archive registration system:

- The registration process 1 2 3 4 5
- Logging in on subsequent visits, once already registered 1 2 3 4 5
- The email instructions 1 2 3 4 5
- Resolution of registration problems 1 2 3 4 5

7. Any other comments relating to the UK Data Archive or the census registration systems:

Your details

Full Name

Department

Email Address

Telephone Number

Which service(s) do you represent?

- MIMAS
- UK Data Archive
- UKBORDERS
- Athens
- CCSR (SARs)
- None of these

If you would prefer your contact details not to be included in the database of respondents to this survey, please tick this box.

The new service will be beta tested in 2002, please indicate below whether or not you would like to be involved in this process.

- Yes please
- No thank you
- Not sure, please contact me again later on in 2002.

© Copyright 2002 University of Essex. All rights reserved. TOP
