

# Update on ONS Proposals for an Integrated Population Statistics System

In October 2003, ONS published two consultation documents. One was on the proposed 2011 Census design ('Census strategic development review: user requirements report' ([http://www.statistics.gov.uk/downloads/theme\\_population/user\\_req.pdf](http://www.statistics.gov.uk/downloads/theme_population/user_req.pdf)) and the other presented ONS's vision for the long-term future of population statistics, an Integrated Population statistics System (IPSS). (See [http://www.statistics.gov.uk/downloads/theme\\_population/ipss.pdf](http://www.statistics.gov.uk/downloads/theme_population/ipss.pdf)). This is an update of progress on the IPSS.

## IPSS vision

Improvements in managing information in government, particularly administrative systems, mean that, with the right legal protection, we could in principle use information already held by government to put together much of what we traditionally obtain via the census. Although this will not be possible for 2011, the IPSS proposals presented a way forward in this direction. The proposals set out ONS's thinking on the statistical issues but not on the wider changes that would also be required to the arrangements for accessing and protecting information. The consultation paper acknowledged that there were both benefits and risks in following a path towards this vision, and set out those that had been initially identified

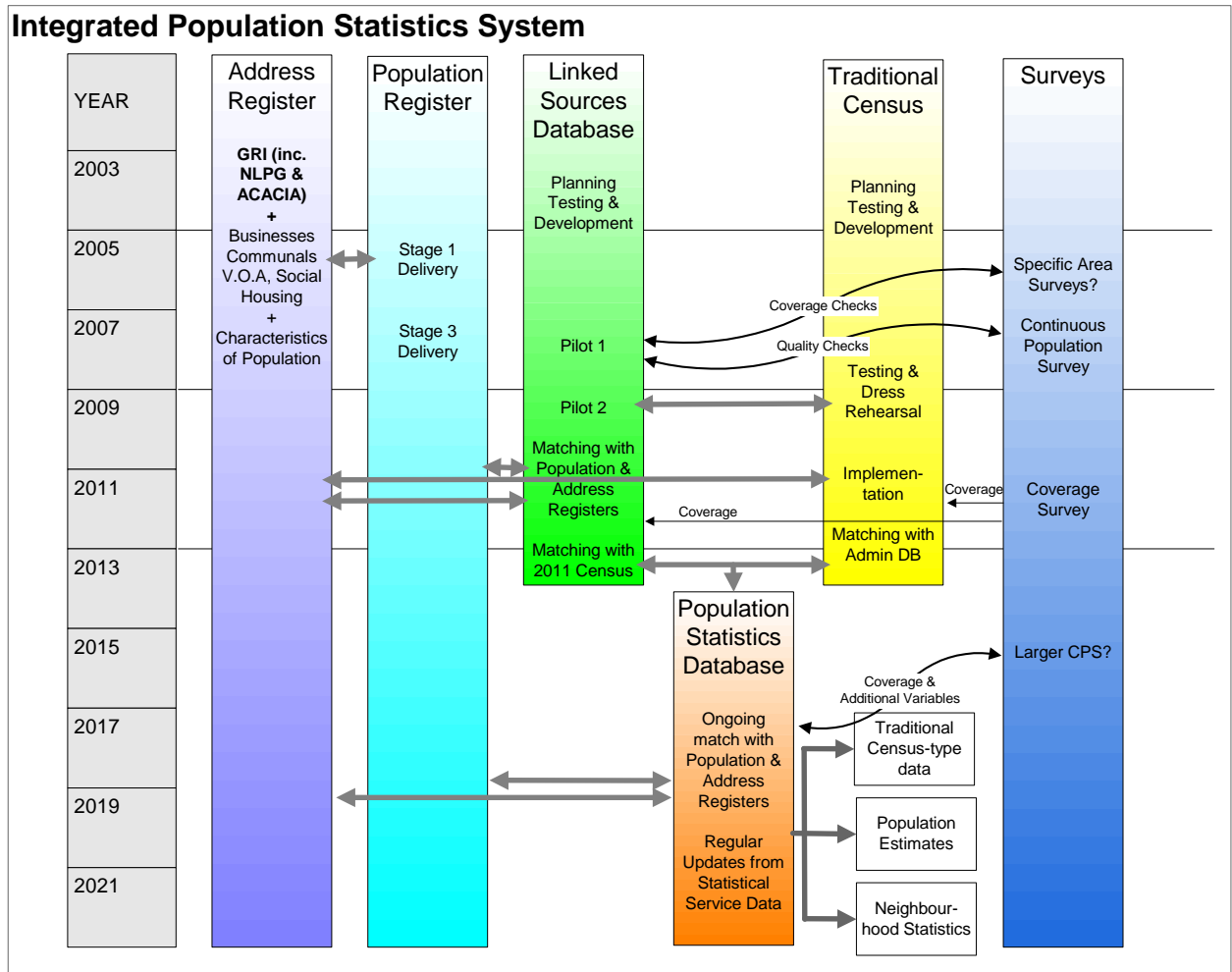
The IPSS vision is to combine census, survey, and administrative data, linked at individual person level, to create a single, comprehensive population statistics database, which is updated over time. This database would underpin all ONS population and social statistics, resulting in significantly improved and more consistent statistics for the government community, the Health Service, academia and the private sector.

The key elements of the proposal are :

- a high quality address register, covering all properties in England and Wales
- a full 2011 census operation
- creation of a linked statistical database, combining administrative and survey data by linking at the individual person and household level;

The aim would be that, following the 2011 Census, the linked statistical and census databases would be combined to create a linked population statistics database. This linked population statistics database would be updated using administrative records, survey data, the address register, and any future population register, with consideration given to how such data answer user requirements. The aim would be for this database, once it was sufficiently populated and updated, to form the basis of all future population statistics produced by the ONS.

**Fig. 1**



**User's comments in response to the IPSS discussion document**

Users were invited to comment on the proposals generally and to identify, from their own perspective, the desirability and scope of such an integrated population statistics system, and any additional risks and benefits. A summary of the comments received can be found in *Annex A*. In general, users generally expressed strong support for the overall proposals for an integrated population statistics system (IPSS), which were described as exciting and visionary. However, nearly all the comments emphasised what a massive undertaking it would be, and a common concern amongst users was that the strategy was high risk and that the proposals did not perhaps fully recognise the risks. Nevertheless, the feeling was that the vision was definitely worth striving for.

**Update on work in progress**

The IPSS is a long-term undertaking and it is inevitable that priority in funding and resource tends to be given to more immediate aspects, such as planning for the 2011 Census and improving population statistics between now and 2011.

Some work has begun on each key statistical element; progress on each is summarised below .

With regard to the wider changes that would also be required in the arrangements for accessing and protecting information, a Government Statistical Service (GSS) Task Force has been looking at issues around data sharing and their report, the GSS Task Force Report on Data Sharing, is due to be published shortly.

With regard to the aim of establishing a single, high quality UK address database, ONS is working with ODPM, other parts of central and local government and other agencies that use addresses.

### **1. A full 2011 census operation**

The ONS is preparing for a 2011 census. The first stage was to review whether there was a continuing need for the information traditionally provided by a census. The consultation document mentioned earlier (Census strategic development review : User requirements report, released at the same time as the IPSS document, concluded that there was a continuing need for this information. A review of alternatives had concluded that no alternative could replace the census by 2011. A high-level design for the 2011 census was published at the end of 2003. During 2004 and 2005 work has been progressing on development of the census business case (including cost and benefit analysis); research and more detailed design for the census; and development of a procurement strategy (for procuring the services to be purchased from the market).

To help achieve improved UK consistency, a 'statement of agreement between the Registrars general' for England & Wales, Scotland, and Northern Ireland was published in February 2005 confirming their aim to work in unison to meet the needs of UK data users.

Agreement has been reached (between the three Registrars General) on what population base will be used for the 2011 census across the UK : we will count usual residents, but also collect some information on visitors on census night so that some outputs can be produced based on usual residents plus visitors.

An initial census questionnaire is required by mid 2006 for the planned 2007 Census test. User consultation on the content of the 2011 Census therefore began on 16 May 2005, with the launch of a paper - '2011 Census: Initial view on content for England and Wales'. This paper invited responses (until 5 August) from all users in support of topics and classifications of interest, using a standard template. The consultation document was supported by a series of six roadshows (attended by over 300 people in total, including local and central government users) across England and Wales. Questionnaire development and further consultation will continue for the next eighteen months.

### **2. Towards a future population register : the Citizen Information Project**

ONS has been running a project - the Citizen Information Project (CIP) - on behalf of the Treasury with the aim of saving public money and improving services to citizens by increasing the sharing of basic citizen contact details (such as name and address, date of birth, etc) across central and local government. A more detailed introduction to the CIP can be found at <http://www.gro.gov.uk/cip/>. The options considered are all based on some kind of population register or registers, for example such as the single 'notification register' that underpins all public administration in Sweden.

In addition to the financial and services issues, CIP has also looked at statistical requirements for a population register, both to underpin population statistics work and more generally. This includes possibilities along the lines of the IPSS vision, bringing together a range of administrative, survey and census data for statistical purposes within a secure statistical environment, for example as now done in Norway, Finland and The Netherlands. The consideration of the statistical requirements will be published along with other CIP material. The CIP recommendations are currently being considered by the Government and the Chief Secretary will make a statement to the House after the summer. The whole CIP report will be published by ONS at the time of the Chief Secretary's statement.

### **3. Creation of a linked statistical database**

Some research has begun, however, initially focusing on the technical challenges and statistical benefits of individual - level linkage with the intention of delivering proof of the IPSS concept . ONS are

discussing with the Department for Work & Pensions and GRO Scotland (GROS) a pilot study to establish whether we can successfully link data for England, Wales and Scotland drawn from different sources including 2001 Census, and DWP Benefits data. The intention would then be to extend the pilot to NHSCR information. The objective would be to establish the statistical benefits of an IPSS-type database to ONS and the broader GSS and to assess the potential benefits of this linkage for ONS, DWP and GROS.

Possible long term benefits include:

- improving resource allocation and service provision
- improving the evidence on which policy is based
- reducing the cost and burden of government. If accurate data could be derived from administrative sources, the number of questions in the Census and other surveys could be reduced (and, potentially, the Census could be dispensed with altogether) with a saving of public funds and a reduction in the burden on the citizen.

Possible short term benefits include:

- enabling DWP to produce statistics on benefit take-up by ethnicity, tenure and education.
- augmenting intercensal population statistics by adding ethnicity to migration statistics and improving household modelling
- assisting 2011 Census by providing a basis for targeting coverage and coverage adjustment.

The pilot would involve the matching of a small (around 1%) sample of data from each source. Ultimately the aim is to deliver enough information to decide whether to proceed further with the IPSS, based on whether we can create an individual-level linked database of sufficient quality and the identified costs and benefits. Discussions continue between ONS, GROS and DWP on the legality and practicality of sharing data for the pilot.

---

## **Annex A**

### ***Proposals for an Integrated Population Statistics System: User's comments***

#### **Background**

The ONS Discussion Paper *Proposals for an Integrated Population Statistics System*<sup>1</sup> set out a vision for a future population statistics system in England and Wales, bringing together data from a number of possible sources to produce a statistics database that would contain person-level information covering the whole population.

The Paper acknowledged that there were both benefits and risks in following a path towards this vision, and set out those that had been initially identified. Users were invited to comment on the proposals generally and to identify, from their own perspective, the desirability and scope of such an integrated population statistics system, and any additional risks and benefits.

This document summarises the many specific comments that were received during the course of the recent consultation exercise and discussed at the joint ONS/RSS Conference in November 2003, and refers particularly to those which were common to a number of users. Many comments were raised in the form of questions posed.

A list of those organisations and individuals who made responses to the proposals for both the Integrated Population Statistics System, and/or the accompanying paper on the proposals for an 2011 Census design<sup>2</sup>, are listed in Annex B. The ONS are most grateful to all those who took the trouble to respond.

## Overall comments

Users generally expressed strong support for the overall proposals for an integrated population statistics system (IPSS), which were described as exciting and visionary. Nearly all the comments emphasised what a massive undertaking it would be, and a common concern amongst users was that the strategy was high risk. Nevertheless, the feeling was that the vision was definitely worth striving for.

Several comments urged ONS to concentrate on producing best quality administrative data in conjunction with local authorities and other statutory organisations if there was to be confidence in the robustness, reliability and validity of the IPSS. It was noted that many local authorities and others are already undertaking linking of some datasets and there were a number of offers to share knowledge and experience.

Many users commented that implementation would be a massive task, in particular in undertaking a standardisation of all survey and administrative data. The harmonisation of definitions of administrative sources data was seen as an area of particular concern. A significant amount of time needs to be built into the project to allow for change to survey and administrative data by government departments, local authorities and others. Also, there would be a need to consider how to deal with sources which do not have the resources or inclination to make the necessary changes.

A common view was that there would be a need to continue collecting information from traditional-style censuses (or, at least, very large-scale surveys) after 2011 in order to recalibrate the system. This applies especially in those subject areas where administrative records do not supply sufficiently up-to-date information and where surveys do not provide complete coverage. Such subject areas include household membership/relationship and characteristics, cohabitation, non-household populations, economic *in*activity, accommodation type, unpaid care, hours worked, level of qualifications, employment status, tenure, number of rooms, travel to work.

It was felt that the ONS would need to consider particularly, the output strategy it would adopt in order to make this information generally available.

ONS acknowledges the technical challenges but one user commented that if they are to be seriously addressed, they will require at least as much technical development effort as was invested in the 2001 One Number Census or Output Area design.

One user commented that there was a lack of explicit commitment to estimating the full population since the proposals, even with enormous political will, offer only linked administrative and survey records at their current quality and do not provide the accurate source of administrative-based population estimates. The sense is given therefore of a new population statistics that does not include all those missed by direct counting methods.

There was one particular comment that the vision was not feasible without a multi-purpose person number as the linkage mechanism, and that this could not exist without a population register.

A common theme amongst the comment was the acknowledgement that the creation of an address list for a population statistics system faces several challenges. It was felt that statistical requirements of a national address list are different to those of other organisations and that it would be difficult for owners of any current lists to accept the validity of a population system which was out of agreement with their own lists. This could lead to tensions between statistical and operational priorities in the maintenance of a national register. It was further considered that the continued failure of the addressing industry to reach agreement even over abstract addressing standards suggests that resolution of the potential conflicts will require high-level government intervention.

There was interest in the concept of 'lawfully resident' and how this was to be defined.

Users noted that IPSS would not take effect until 2013 but were keen to see population estimates from a register coming on-line before then.

The proposals suggest that with explicit and special legal authority such data may be available for statistical purposes. Users were interested to know what ONS envisage as the scope for such uses and the process for use of the data.

Many users were interested to know who would have access to individual level data and for what statistical purposes.

There was interest also in how trans-national commuting will be handled.

## **Benefits**

The benefits identified in the Discussion Paper were generally acknowledged by most users. Specifically:

- Users felt that it was good to see that the database would provide an individual level sampling frame for social surveys. Many would benefit from a consistent and up-to-date sampling frame.
- Input and, possibly, resource from all the several sources referred to in the Discussion Paper could be utilised to ensure that the system would meet a wide range of user requirements as fully as possible and deliver maximum benefits.

## **Additional benefits**

Some additional benefits were identified by users:-

- Outputs will allow much more frequent and comprehensive access to geodemographic data for market modelling and targeting purposes.
- The likelihood of significantly reduced costs for traditional census taking from 2021 onwards.
- Address Register will overcome many of the limitations associated with the PAF and remove the risk of PAF availability being compromised by any future changes at Royal Mail.
- Major benefits to geographical referencing providing it was accessible to all user organisations.
- Timeliness providing capacity to produce continually updated official population statistics.
- If appropriate outputs were conceived, capacity for analysis of population change would be enormously increased and ability to devise resource allocation and targeting schemes that reflect contemporary social conditions much enhanced.
- Potential attribute breadth of the new system would allow new multivariate investigations.
- Resilience to enumeration non-response, which international experience suggests will continue to increase and therefore hamper all future efforts at conventional census enumeration.

## **Risks**

Users agreed with the risks already highlighted in the paper. In particular they noted that:

- Key to the proposal is the reliance on the ability to link various administrative databases.

- One of the weaknesses of the system is the individual-level linkage. Issues of confidentiality must be addressed if it is not to become a white elephant or a black hole into which data is absorbed but from which few useful outputs are observed.
- Public acceptability – there would need to be absolute guarantees that data will not be misused, but even so many people might refuse to complete forms if they knew that their personal data, provided in confidence for a particular purpose could be linked via a unique reference number to personal data provided for other purposes to other organisations.
- Linking of data is a huge undertaking and fraught with difficulties similar to, if not worse than, those experienced in the 2001 Census with regard to tracking people who are constantly on the move, those with more than one address, those living in different households, etc. Keeping track of changes to name and address will prove similarly difficult.
- A severe risk would be in the funding, both long term and short term. Funding in the timescale envisaged would have to be approved through at least three Parliaments.
- Rough sleepers, travellers, refugees, homeless and groups with no permanent address were at risk of being missed, and therefore risk a growing divergence of the IPSS population from the ‘true’ population if such groups were missed systematically;
- The history of the design and implementation of massive computer systems in government and NHS does not inspire confidence.
- Address matching is a complex area.
- There would be a devaluation of the data by disclosure control precautions.
- Many of the proposed datasets do not exist but the system’s effectiveness relies heavily on them all being in place and being reliable systems in their own right.
- Harmonisation of definitions of administrative sources data may be difficult to achieve.
- Aspects of data protection legislation require data held about individuals to be accurate. But measures to prevent disclosure may ‘corrupt’ an individual’s information and produce inaccurate information about that person. This may occur inadvertently with any failure to link individual characteristics correctly. There were concerns about how the accuracy of the data held on each individual would be maintained and verified.
- There were concerns about the currency of datasets in disparate administrative sources, and about the delays there might be in extracting records. It was felt that there might be significant mismatches between data sets through lack of synchronicity.
- The notion that changes of address will be notified sufficiently accurately to allow a picture as accurate as a census count is regarded as highly improbable. It was suggested that evidence is needed from other countries as to the accuracy of population estimates of registers of the proposed type.
- The proposals would put demography and planning on to a wholly different footing, one that is less democratic because of its exclusion of parts of the population living in the UK.
- Some users thought that the proposals appear to be based on over-optimistic timescales for approval to create such databases taking into account the necessary public debate, legislative changes and great political obstacles.

## **Scope**

Additionally respondents commented on the possible scope and application of the proposed system.

- There were a number of comments relating to the proposal to cover only 'lawfully resident' population and how this is to be defined.
- Many users held the view that it was essential to co-ordinate implementation across government, health service, academia and private sector. The need for co-ordination across the UK was also expressed. A number of users were of the opinion that it was absolutely essential that such a system should be UK-wide otherwise the system would be unable to record migration across borders and would rapidly become unworkable. Users with a GB interest were also keen that maximum read across from E/W proposals applies to Scotland.
- Strong support for proposals for population counts at a disaggregated level, although the release-level geography was not specified in the document.
- Users suggested establishing a high level Advisory Group crossing all sectors to co-ordinate the general implementation of the system. More focused groups could then be formed for specific areas of development as required.
- Data from different sources should be drawn in as widely as possible to ensure the robustness and continuing usability of data. Vital to achieve through joined-up thinking between the many government departments who collect this data and the development of public trust for the ways in which these data will be used.
- Key issues would be level of access to data which has not been heavily treated for disclosure at very small area level

## **Output**

- Some users are looking for frequent population estimates at OA level on broad socio-economic groups/characteristics
- Main advantage of Census is ability to cross-tabulate at small area level. IPSS must be able to do the same.
- Several users commented that there are some Census datasets for which there is no equivalent administrative data – e.g. travel to work at the small area level.

## ***Annex B***

### ***List of respondents***

#### **Organisations**

Bath and NE Somerset; Bristol City; North Somerset; South Gloucestershire; Joint Strategic Planning and Transportation Unit (joint submission)

Blaenau-Gwent Local Authority

Caerphilly Local Authority

Cheshire County Council

Central and Local Government Information Partnership (CLIP) Census Sub Group

Centre for Census and Survey Research (CCSR) - Ludi Simpson

Demographics User Group

Department for Work and Pensions (DWP)

Department for Rural Affairs (DEFRA)

Department for Transport (DfT)

Derby City Council

Disability Rights Commission

Economic and Social Research Council - Ian Diamond

Economic and Social Research Council (ESRC)/Joint Information Systems Committee (JISC) 2001 Census Programme - David Martin, Co-ordinator

General Register Office Scotland

Greater London Authority

Help the Aged

Herefordshire Council

Hertfordshire Census Users Group (Hertfordshire Police, Public Health Intelligence and Hertfordshire Health Informatics service representing the Primary Care Trusts and Strategic Health Authority, Learning and Skills Council and Hertfordshire University).

Home Office

Institute for Social and Economic Research

Leicestershire County Council

Liaison Group for Population Statistics (LA Side)

Local Government Data Unit – Wales

Luton Borough Council

Manchester City Council

Market Research Society - Census Geodemographics Group

Medway Council

National Centre for Social Research

National Housing Federation

Royal National Institute for the Blind (RNIB)

Stoke on Trent City Council

Suffolk County Council

Tees Valley Local Authorities – (Darlington, Hartlepool, Middlesbrough, Redcar and Cleveland, Stockton-on-Tees)

Welsh Assembly Research Network

Welsh Assembly Government – Statistical Directorate

**Private individuals**

Philip Redfern

Stefan Webb

*The ONS are most grateful to all those to took the trouble to respond.*