

CENSUS ADVISORY GROUP

Advisory Group Paper (08) 04 - Addressing for the 2011 Census

1 Introduction

This paper presents:

- an overview of the 2011 Census addressing strategy;
- the scope of the pilot programme being initiated; and,
- where Census is seeking LA involvement in the strategy and the pilot.

Advisory Group members are asked to:

- provide views on the strategy and pilot project; and,
- support the strategy and pilot through communication within and with their LAs.

Shift in focus

The addressing strategy, attached at Annex A, has changed focus from paper AG (07) 08 presented to Advisory Groups in Autumn 07. The shift in focus provides a strategy that will deliver an Address Register with better use of ONS resources whilst providing benefits to Address Suppliers, in particular LAs.

The main changes are:

- Full use of the primary sources of address lists, namely the Royal Mail Postcode Address File (PAF), the OS MasterMap Address Layer 2 (AL2), and the National Land and Property Gazetteer (NLPG);
- Local authority participation and endorsement of the Census Address list for their area, providing wide assurance on the quality of addresses underpinning the Census and hence the outputs;
- Subject to approval, sharing of addressing intelligence gained through development and address checking with address suppliers improving the overall quality of addressing used throughout England & Wales.

2 Addressing Strategy

The focus of the addressing strategy attached at Annex A is to develop an address register (AR) that meets Census requirements in terms of quality and coverage but also an address register that LAs are confident will effectively underpin Census enumeration and outputs.

The strategy achieves this through:

- Optimal use of existing addressing products;
- Engagement with LAs to utilise expertise within the LAs to enhance the quality of the AR for their area ;
- Seek LA endorsement of the Census AR;
- A targeted address checking exercise before Census day in areas where the products have high levels of discrepancies/unmatched records and in areas of high multi-occupation.

Royal Mail, Ordnance Survey, IDeA and the LGA have seen the strategy and are supporting ONS through regular data supply and advice throughout the period of the pilot.

3. LA Involvement

Census is seeking to work with LAs to successfully implement the addressing strategy and deliver a fit for purpose address register for the Census. In particular, by:

- Working in partnership with LAs to design and implement a process that is likely to deliver the best possible AR nationally.
- Working with LAs to encourage and support improvements in their LLPGs ahead of the verification of addresses for the 2011 Census Address Register.
- Identifying mismatches between address sources and working with LAs to develop tools and supporting information that facilitates the most effective resolution of mismatches with LAs.
- Working with LAs (and other suppliers of addressing information) to understand the process for updating and maintaining the address sources in order to obtain updates to the AR near to Census Day.
- Working with LAs to develop the process for LA endorsement of the Census AR to ensure that the process is as efficient as possible and supports LAs in the endorsement of the Census AR for their area.

4. Pilot – 24 LAs

A pilot exercise is being conducted in the middle of 2008 to test and refine the process set out in the addressing strategy. The pilot will:

- Develop an efficient process for cross-matching address sources to identify address mismatches to be investigated
- Assess address data sources against evaluation criteria, e.g. time-series, positional accuracy, multi-occupancy

- Improve our understanding of the main address sources and how best to anticipate update information near the time of the Census to get an AR as up-to-date as possible.
- Improve understanding of variability between LLPG sources.
- Examine other sources of data, including extra data held by LAs, to assess their relevance, particularly for difficult enumeration addresses in 'hard to count' areas and communal establishments.
- be used to develop the timeline and process for LAs 'endorsing' their address lists ahead of the 2011 Census.
- derive and develop methods and information to prioritise/identify areas for Census Address checking, with a focus on difficult areas including those with high multiple occupancy.
- assess the risks associated with this strategy and develop mitigation plans to minimise the risk.

A list of LAs selected to participate in the pilot exercise is attached at Annex B.

The main events within the pilot are:

Description	Dates
Letter to Pilot LA Chief Executives introducing the pilot	March
Regional workshops with LLPG custodians	April
LAs desk and field work to verify mismatches between sources	May - July
ONS field work in sample of areas	May - July
LAs provide resolved mismatches to ONS	August
ONS report on address source characteristics and change profile	September
Workshop to review pilot with LLPG custodians	September
Report on key findings	November

5. Supporting activities

In addition to the pilot exercise, there are a number of key streams of activities to support the overall development of the strategy. These include:

Licensing/Costs

This will involve establishing the terms and conditions of acquiring the address products and any incurred costs for Census use. It will also be necessary to agree the basis on which information about the sources can be shared with the other suppliers.

Welsh language

Research will be carried out into the availability of addresses in Welsh within the address products or from other sources.

Software evaluation

Assess software products that will enable the address products (full supply and change only updates) to be loaded into the ONSG Oracle and SQL Server database.

Communal establishments

The suitability of other products containing specific communal establishment types will be investigated to assess if they provide extra addresses and add value (e.g. capacity information such as number of residents, or contact details) to the address registers. These extra data suppliers will be dedicated and respected authorities on each of these particular types of communal establishments

Address data representing the following types of communal establishments will be evaluated:

- University halls of residences and boarding schools
- Sheltered and residential care/retirement accommodation
- Hospitals and other medical establishments with in-patient facilities
- Camp and caravanning sites
- Hotels
- Bed & breakfasts and guest houses
- Police and fire stations
- Prisons and detention facilities (e.g. Young Offenders Institutes, Immigration Removal Centres Secure Childrens Homes, Probation / Bail Hostel)
- Foreign and national military bases

National programme of communication

- Initiate a programme of communication with key stakeholders, namely LAs and address product suppliers to promote and seek support for the strategy.
- An Address Steering group will also be established with very senior stakeholder and supplier involvement to establish principles and steer the process through the Census.

6. Next steps

The pilot exercise is planned to conclude in October 2008, with the main interaction with LAs in April to September. An update on progress and a summary of the key findings from the pilot will be presented to the next round of Advisory Groups in Autumn '08.



2011 Census Addressing Strategy

1 Introduction

This addressing strategy for the 2011 Census identifies the key areas of development work to be undertaken to achieve the 2011 Census addressing objectives:

- To develop an Address Register that maximises coverage and quality to support critical parts of the Census operation: questionnaire delivery (mostly post-out with a small amount of hand delivery), follow-up and the production of outputs. All of which will be needed to ensure that the Census meets its strategic objective of maximising response and minimising differential non-response.
- To get buy-in from stakeholders to the register we use, particularly from LAs.
- To achieve the best possible address register in a cost effective manner.

2 Background

Without a single definitive national address register for England and Wales, Census needs to assess the competing address products, Ordnance Survey's MasterMap Address Layer 2 (AL2), the National Land and Property Gazetteer (NLPG) and the RM Postcode Address File (PAF) for the 2011 Census. Other supplementary address data sources are needed for communal establishment and special enumeration addresses. The products are being assessed against a range of criteria in the Census Household Frame requirements. Phase 1 of the research focused on the coverage of AL2 and NLPG in the Census Test areas during the address checking exercise in September/October 2006. Phase 1 of the research (<http://www.statistics.gov.uk/census/2011census/news/default.asp>) concluded that both products fell short of census requirements and further work was required to understand the two products and how they could be used to develop an Address Register for the Census.

3 Addressing Strategy

There are four main strands in the addressing strategy which aims to develop an Address Register of sufficient quality to support the aims and objectives of the Census:

- verification of addresses:
- Engage with LAs about the addresses, seek their support in address resolution and seek their 'endorsement' of the address list developed for their area;
- ONS conducts an Address Check before Census day to supplement the address list with households and addresses not already identified on existing products, based on a prioritised list of areas for fieldwork to allow flexibility in workloads, with respect to quality and cost targets.
- subject to approval, share the addressing intelligence gained during the first three phases with Address product suppliers.

3.1 Verification of addresses

The objective is to maximise the coverage and quality of the Census address register used for enumeration and the production of outputs. This will be achieved through matching the three main address products (RM Postcode Address File (PAF), OS AL2 and NLPG), verification of unmatched addresses by LAs, and then supplementing the addresses matched and verified with Council Tax and Non-domestic rates addresses from VOA, Utility addresses, and Multi-residence addresses from RM.

3.1.1. The matching and verification of addresses will be a four stage process consisting of:

Primary matching

- Link to the OS product to add grid references and building footprints to the PAF. This will be a very high-level of matching because there is a key on the OS and RM products that facilitates matching.
- Match the NLPG to the PAF to add addresses that exist on the NLPG that are not on the PAF and to spot PAF addresses not on NLPG. Matching accuracy between these two products is expected to be lower as the NLPG does not use the PAF as a source product and is, on the whole, in a different address structure.

Secondary matching

- A second round of matching will be undertaken for those NLPG and PAF addresses that don't match each other. This second round of matching will use, largely established, 'fuzzy' matching methods to match addresses where the addresses are similar but in a slightly different format/spelling and wouldn't be matched at direct matching. Geographic intelligence (buildings, features and roads) will be used to support the matching process.

Computer assisted manual matching

- All addresses still not matched after the second round will be subjected to some level of manual matching, possibly using a service bureau that is geared up for such work. This would involve people to compare the two addresses and using supplementary information (gazetteers, Web, yellow pages, MasterMap plots) to confirm, or not, matched addresses.

Verification of unmatched records

- All remaining unmatched addresses will be sent to local authorities (in a well structured form) to verify the addresses. This includes addresses on the PAF and not on the NLPG and addresses on the NLPG but not on the PAF. The LA will verify each address in the office or in the field (sometimes with ONS also involved), which will be passed back to ONS and incorporated on the base address list. RM and OS would also be informed of any changes.
- As the majority of discrepancies are thought to be long-standing problems in Address Products, the verification process will be done on a rolling basis with all LAs at least once before Census day, with the majority of the verification likely to be done in 2009.
- After the LAs have verified the unmatched addresses, they will be asked to 'endorse' the address list for their LA. In addition to the list, the LA will receive sufficient information and training to make the verification effective and efficient. Para 3.2 has further details about LA endorsement.

In addition to the matching and verification of addresses noted above, the creation of the base address list will use information from other data sources to ensure that all of the information required for the Census address register is timely and accurate as possible. This is likely to include:

- information on Communal establishments (CEs) which would be derived from additional data sources which are known to be better at recording CEs.
- information from VOA, TV licenses, utility suppliers and others would be used to identify any additional addresses not already created through the matching of PAF, OS and NLPG.
- information on multi-occupancy drawn from RM and potentially other sources.

3.2 LA endorsement

Each LA will be asked to 'endorse' the address list for their LA. The 'endorsement' process will:

- assure the LAs that the address list used for the Census is as accurate and timely as possible; and,
- assure the LAs that the address list is consistent with their NLPG records providing a sufficient level of confidence to LAs and ONS in the Census design and the delivery of accurate population estimates.

To commit to 'endorsement', the LAs will require assurance on two key points:

- verification of unmatched addresses. They will undertake this verification themselves, see section 3.1.
- assurance of the process for updating the address register between sign-off and Census Day. For example the addition of updates to the PAF, NLPG plus new-build intelligence and the incorporation of information found at address checking.

3.3 Address checking by ONS

ONS will conduct an address checking exercise in areas where the products have high levels of discrepancies/unmatched records and in areas of high multi-occupation. The prioritisation model and the area choice thresholds to be used will form an important element of the next stage of research. In addition, some further random checking of LA work may be required, and it may be necessary to do more extensive work in areas where the LAs are less able or less responsive.

3.4 Sharing address intelligence

The processes noted at 3.1 and 3.3 will provide a significant amount of address intelligence. The strategy is to share the address intelligence gained during the address verification and matching, and the address checking stages with the Address product suppliers. The sharing of this information will enable the address products suppliers to update their products to deliver an addressing product with improved accuracy and coverage to their users. The sharing of address intelligence is subject to the approval of ONS and the Address product suppliers: RM, OS and IDeA (custodians of the NLPG).

4 Fulfilling Addressing objectives

A programme of research, phase 2, has been initiated to develop and inform the process for achieving our addressing objectives. Critical to this research is a supplementary stream of work to engage LAs and other stakeholders (address product suppliers and other important interested parties), to ensure we maximise the accuracy of the address list and to ensure buy-in to the address list used for the Census. The objectives and scope of these two streams is set out in paras 4.1 to 4.7.

Phase 2 Objectives

- Develop an efficient process for the implementation of address verification through evaluation against specific criteria, e.g. time series, positional accuracy, multi-occupation etc.
- To develop the timeline and process for LAs 'approving' their address lists ahead of the 2011 Census.
- To achieve LA support to the process of developing an Address list for the Census.
- To derive and develop methods and information to prioritise/identify areas for Census Address checking.
- To assess and the risks associated with this strategy and develop mitigation plans to minimise the risk.
- To gain approval for the sharing of address intelligence gained during address verification and address checking.

The intelligence gained during the phase 2 research will be used to develop and implement a set of processes to construct an Address Register for the 2009 Rehearsal.

4.1 Address verification

Matching

- Develop a methodology for maximising the automatic and secondary matching of address data sources, primarily PAF, OS, NLPG, plus manual resolution of mismatches.
- Assess the volumes of automatic matching (primary and secondary) and the system requirements, and cost, necessary to support this level of matching on a national basis
- Assess the unmatched addresses and whether undertaking a manual matching exercise delivers significant benefits compared with the cost.

LA verification

- Secure LA buy-in to the process of verification as a first step to achieving LA buy-in to the sign-off strategy for 2011;
- Assess the quality and timeliness of the LA verification process to inform the development, and timing, of the verification process to be carried out during 2009 and 2010.

ONS field test

- ONS selects a sample of areas from the pilot LAs and does an on the ground check. The field check will provide a true' dataset with which to assess the quality of the address products, particularly how many addresses don't exist on any of the key products; and to assess the quality of the LA verification.

Assessment of results:

Using the information gained out of the matching exercise, the LA verification, the ONS field check and any other supplementary sources assess the level of completeness of the resulting address lists to inform:

- the development of the LA sign-off process, particularly the amount of verification likely to be required; and,
- the processes and work needed to find multi-occupied addresses.

4.2 Address churn and dynamics

- Identify and understand the data sources and processes used to create the three main addressing products and the supplementary address sources and the associated quality issues.
- Track and quantify the timeliness of the three main addressing products to assess the amount and nature of turn-over. This will be particularly important to:
 - assess when final updates can be made to the Address Register before Census Day and the volume of updates and resource required to undertake the update process;
 - finalise the Address Register design which will be used to assure LAs during the sign-off that the Register for their area is up to date and late updates to address products will be incorporated before Census day; and,
 - understand how best to intercept the updates being made to the address sources at any point in time.

4.3 Address Check area selection

- Develop a model to facilitate the selection of small areas for address checking. This model would consider various data from 2001 Census, administrative and commercial sources to rank areas where the quality of the address lists isn't acceptable and would benefit from address checking, in particular areas of high multi-occupancy.

4.4 LA support and buy-in

- Develop a strategy for working with LAs to implement the addressing strategy, their involvement in the verification of unmatched addresses and the 'endorsement' process. This will include a regime of very senior involvement to establish principles and steer the process.

4.5 LA endorsement

- Develop a process and timeline, based on the findings from the phase 2 research, for the LA 'endorsement' process. This will have some overlap with the LA support and obtaining assistance to the development of the process.

4.6 Strategy for engagement with key stakeholders

- Develop a strategy for engaging with key stakeholders, namely address data providers, to ensure that they understand the Census addressing strategy and our intentions to work in partnership to develop the best address register. Other stakeholders that might have an influence on a successful address register will also be identified and targeted.

4.7 Assessing risk

The phase 2 research will assess the risks associated with adopting the strategy outlined in section 2 and develop mitigation plans to minimise the risks. The main risks already identified and which will be investigated as part of the phase 2 research are:

- Volume of matched addresses might be lower than expected.
- Volume and cost of matching may be impractical and prohibitive;
- Volume of unmatched addresses sent to LAs for verification may be prohibitive;
- Some LAs may not have the resources to verify unmatched addresses;
- Levels of LA engagement and consistency are highly variable; and
- Use of PAF and the extra licensing costs that this is likely to attract.

5 Key Deliverables and milestones

An indicative list of key deliverables for delivering the address strategy are set out below. The list is likely to move and change as the process develops and as more information is gathered.

Deliverables/Milestones	Date
Letter to pilot LAs selected for address verification	March 2008
Agree address sources and matching methods	April 2008
A high level strategy for engaging with LAs	April 2008
Launch of national campaign for addressing with LAs	May 2008
List of unmatched addresses to pilot LAs for verification	May 2008
Prioritisation model (first iteration) for fieldwork complete	September 2008
Verification of Addresses by pilot LAs and ONS complete	September 2008
Report on Address churn and dynamics	September 2008
Plan for 'rolling verification'	September 2008
Communication with LAs about rolling verification	November 2008
Verification of Addresses by Rehearsal LAs and ONS complete	May 2009
LAs endorse Rehearsal address lists	June 2009
Prioritisation model (second iteration) for field work complete	October 2009
Census rehearsal	October 2009
LAs signed up to endorsement protocol	March 2010
First supply of address products for matching for 2011 Census (Address Base Day)	March 2010
LA endorse Census address lists	October 2010
First addresses to Route A supplier for overprinting	November 2010
Supplementary addresses update (e.g. new builds) to Route A supplier	March 2011
Census Day	March 2011

ANNEX B

LOCAL AUTHORITIES SELECTED FOR THE 2008 ADDRESSING PILOT

lad_id	LA Name	In Wales
00AW	Kensington and Chelsea	
00BE	Southwark	
00BG	Tower Hamlets	
00BM	Bury	
00CJ	Newcastle upon Tyne	
00HB	Bristol	
00NC	Gwynedd	Yes
00NQ	Ceredigion	Yes
00NS	Pembrokeshire	Yes
15UE	North Cornwall	
16UF	Eden	
16UG	South Lakeland	
17UK	South Derbyshire	
18UL	West Devon	
19UC	Christchurch	
19UJ	Weymouth and Portland	
21UD	Hastings	
23UC	Cotswold	
24UD	Eastleigh	
26UB	Broxbourne	
30UF	Fylde	
33UD	Great Yarmouth	
44UE	Stratford-on-Avon	
44UF	Warwick	