2011 Census Address Register Development

Garnett Compton

2 October 2008
2011 Census – address register context

• Key Census design changes
  – Questionnaire tracking
  – Post out majority of questionnaires
  – Flexible application of field force
  – Key data source for QA of Census estimates

• Census aim for address register
  – Achieve best possible coverage from all lists
  – Key stakeholders confident in AR
  – Best value for money

• Result will be a better basis for Census but **NOT** a list for general public use
Address register development process

- ONS matches 3 national address files
  - Royal Mail - **PAF**
  - IDeA - **NLPG**
  - Ordnance Survey – **AL2**
- Matching based on address reference, text and spatial relations
- Mismatches sent to IDeA, RM and OS for resolution
- Remaining mismatches sent to LAs for resolution
- ONS checks addresses on ground in c. 30% of areas where most uncertainty remains
  - Also looking for missing addresses and multiple occupation
- Ongoing change updates from OS and IDeA
- Currently piloting with RM, OS, IDeA and 24 LAs
Address matching process

- combination of data linking, textual comparison and testing of spatial relations
- each matching algorithm is applied independently
- a confidence level is generated from the different and combined match results
- resulting residue of addresses with low match confidence are prioritised according to type of mismatch
- unresolved addresses are provided to national address data suppliers and LAs to help resolve
# Priority of unresolved addresses

The **Priority indicator** is related to the type of address-matching and confidence level achieved by ONS.

<table>
<thead>
<tr>
<th>Priority</th>
<th>Description</th>
<th>Form of match</th>
</tr>
</thead>
</table>
| **10-12** | High Priority. LA requested to make every effort to check address. | • Complex – Purely Residential  
• Complex – Mixed Commercial and Residential  
• Unmatched addresses outside of buildings |
| **6-9** | Address should be validated by LA if resources permit | • Flat mismatches  
• Addresses from one product not matching addresses from other products  
• Buildings Missing Addresses |
| **1-5** | Validation not required. Not sent to Suppliers / LAs | • Classification Mismatches  
• Many-to-One relationships  
• Missing New Entries  
• Purely Non Residential  
• Address matched with low confidence |
## Category 9 Anomaly – Flat Mismatch

<table>
<thead>
<tr>
<th>Anomaly Key</th>
<th>BS7666_SAON</th>
<th>BS7666_PAON</th>
<th>STREET</th>
</tr>
</thead>
<tbody>
<tr>
<td>mgeo:3000000011</td>
<td>GROUND FLOOR AND FIRST FLOOR</td>
<td>10</td>
<td>GROSVENOR GARDENS</td>
</tr>
<tr>
<td>mgeo:3000000011</td>
<td>SECOND FLOOR FLAT</td>
<td>10</td>
<td>GROSVENOR GARDENS</td>
</tr>
<tr>
<td>mgeo:3000000011</td>
<td>THIRD FLOOR FLAT</td>
<td>10</td>
<td>GROSVENOR GARDENS</td>
</tr>
</tbody>
</table>

### NLPG

<table>
<thead>
<tr>
<th>Anomaly Key</th>
<th>Sub_Building</th>
<th>Building_Number</th>
<th>Street_Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>mgeo:3000000011</td>
<td>FLAT 2</td>
<td>10</td>
<td>GROSVENOR GARDENS</td>
</tr>
<tr>
<td>mgeo:3000000011</td>
<td>FLAT 3</td>
<td>10</td>
<td>GROSVENOR GARDENS</td>
</tr>
<tr>
<td>mgeo:3000000011</td>
<td>FLAT 4</td>
<td>10</td>
<td>GROSVENOR GARDENS</td>
</tr>
<tr>
<td>mgeo:3000000011</td>
<td></td>
<td>10A</td>
<td>GROSVENOR GARDENS</td>
</tr>
</tbody>
</table>
Results from address matching

- Projected target results presented in July 08
- ‘First cut’ results achieved by August 08
- ‘Second cut’ results achieved by September 08

<table>
<thead>
<tr>
<th>Matching algorithm</th>
<th>Projected</th>
<th>First Cut</th>
<th>Sec Cut</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference match</td>
<td>72 %</td>
<td>72 %</td>
<td>72 %</td>
</tr>
<tr>
<td>Textual match</td>
<td>87 %</td>
<td>87 %</td>
<td>88 %</td>
</tr>
<tr>
<td>Spatial match</td>
<td>89 %</td>
<td>89 %</td>
<td>90 %</td>
</tr>
<tr>
<td>Unmatched addresses resolved by additional Textual and Bureau match by ONS</td>
<td>94 %</td>
<td>&lt;96 %</td>
<td>&lt;97 %</td>
</tr>
<tr>
<td>Unmatched addresses resolved additional matching by suppliers</td>
<td>95 %</td>
<td>&lt;97%</td>
<td>&lt;99.5%</td>
</tr>
<tr>
<td>Unresolved address residual to LAs</td>
<td>5 %</td>
<td>&lt;3%</td>
<td>&lt;1%</td>
</tr>
</tbody>
</table>
Map of unresolved addresses
Addressing pilot - key lessons learnt (so far)

- Frequent synchronisation of address data from suppliers – from monthly to either weekly or daily
- The importance of VOA and alias addresses in the matching process
- Avoiding duplication of work already done by other parties
- Improve LA boundaries in PAF (postcode geog ≠admin)
  - Estimate this will reduce unresolved addresses by 0.5 %
- Incorporate matching processes of address data suppliers into ONS methodology
- Improved communication with LAs
  - support for LA resolution from top of the office needed
- New process of change management needs development
  - to bring in resolved and new addresses in a managed way
Going Forward

• Address matching has proved to be very successful

• The real challenges going forward are:
  – Capturing New Addresses
  – Identify Multiple Occupancy

• Gaining support from key stakeholders, particularly LAs.

• Developing a slick and simple process for change management and communication