Virtual Microdata Laboratory

Access to Confidential Data

Richard Welpton
richard.welpton@ons.gsi.gov.uk
Summary

• Background
• Our place in the UK
• How it works
• Governance
• Managing researchers
Background

• VML established 2004 to provide secure research access to ONS business data
• Original aim to support research into business activity
• 2008: Majority of applications related to studies of earnings
• 2009: Support for Migrant Worker Scan (Titchfield)
• **Future**: Default secure location for sensitive data across Government
• **Future**: Continued expansion across GSI
• **Future**: Social Data
Our place in the UK

More confidential, more secure

No release → Virtual microdata laboratory → Special licence → UKDA → Web

Less confidential, easier access

Not anonymised

Business data, Census data

Census, health data, OGD access to business data

GHS LFS

Aggregate data

Aggregate data
VML Setting

- **Secure lab**
  - Data read only
  - Work area read-write

- **Researcher**
  - Logs in to server
  - Only screen shots, no data

- **VML staff**
  - Direct access to server

- **Automatic transfer**
  - Archiving of all transfers
Governance (1): Access to the VML

1: Apply for Approved Researcher status
2: Application checked by VML staff
3: ONS Microdata Release Panel
   - Project has a “valid statistical purpose”
   - Researcher is “fit and proper”
4: Researcher(s) are notified of approval
5: Training seminar and VML induction compulsory
   - Legal background
   - Statistical Disclosure Control training
6: Researcher free to use VML facilities
Governance (2): VML security model

- valid statistical purpose
- trusted researchers
- anonymisation of data
- technical controls around data
- disclosure control of results

⇒ safe use

safe projects + safe people + safe data + safe setting + safe outputs
Governance (3): Security

- VML is a secure ‘box’
- All statistical outputs are reviewed by VML
- Person risk: SRSA introduces strong penalties
- Independent security review:
  The approach adopted is commendable and considerably exceeds similar practices in Government and industry
- We hold both ONS and OGD data – for linking
- Not 100% safe – 2 breaches from 450 researchers over 4 years
- A model for secure/efficient data access within the UK and abroad
- Reputation as a secure and convenient store for data
## Managing Researchers (1)

<table>
<thead>
<tr>
<th>Data management: researcher as risk</th>
<th>Researcher Management: Researcher as collaborator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explaining security policy</td>
<td>‘we’re doing this to protect the data’ (from you)</td>
</tr>
<tr>
<td>Limiting quantity of results</td>
<td>‘you must limit your output to reduce the chance of disclosure’</td>
</tr>
</tbody>
</table>

Managing Researchers (2): Why manage?

Benefits (of being nice!):
• Increased communication
• Increased understanding
• Increased cooperation
• Effective change management
• Better data security
• Better research
• More efficient use of NSI resources