



Office for  
National Statistics

# Virtual Microdata Laboratory

**Access to Confidential Data**

Richard Welpton

[richard.welpton@ons.gsi.gov.uk](mailto: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



Not  
anonymised

*Business data,  
Census data*

*Census,  
health data,  
OGD access to  
business data*

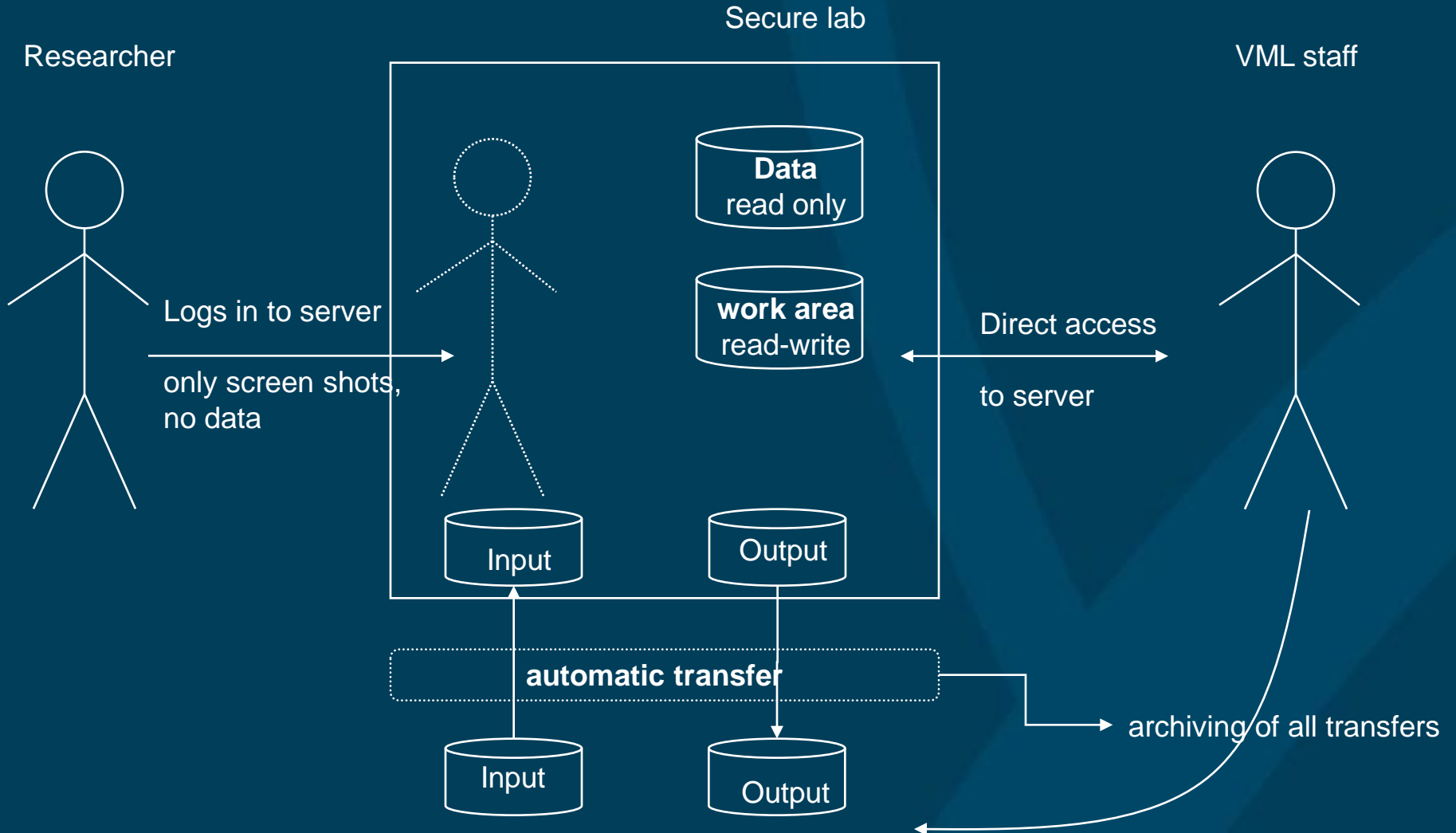
*GHS  
LFS*

*Aggregate  
data*

Less confidential, easier access



# VML Setting



# 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                                  safe projects
  - trusted researchers    + safe people
  - anonymisation of data    + safe data
  - technical controls around data                                  + safe setting
  - disclosure control of results                                   + safe outputs
- ⇒ safe use**

# 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)

	Data management: <i>researcher as risk</i>	Researcher Management: <i>Researcher as collaborator</i>
Explaining security policy	<b>'we're doing this to protect the data' (from you)</b>	<b>'doing this allows us to supply you with more detailed data'</b>
Limiting quantity of results	<b>'you must limit your output to reduce the chance of disclosure'</b>	<b>'limit your output because we have finite resources: people who produce good output get their results back quicker'</b>

Source: Desai, T. and Ritchie, F (2010) "Effective Researcher Management"

# 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