

---

# Administrative Data: Problems and Benefits

Matthew Woollard

Director, UK Data Archive / UK Data Service

Facing the Future: European Research  
Infrastructure for Humanities and Social  
Sciences, Berlin, 22 November 2013

---

UK Data Service

---



---

# Overview

- Background
- Benefits and Problems
- Administrative Data Task Force
- Administrative Data Research Network
- Implementation Challenges



---

# Background – Administrative Data

## **What is administrative data?**

Administrative data refers to information collected primarily for administrative (not research) purposes.

This type of data is collected by government departments and other organisations for the purposes of registration, transaction and record keeping, usually during the delivery of a service.

[Source: ADLS website]

---

# Key British examples

- Department for Work and Pensions Customer Information Systems
- Her Majesty's Revenue and Customs Customer Information Systems
- Offenders Index / Police National Computer (MoJ)
- English/Welsh School Census (DfE)
- National Pupil Database (NPD)
- Student Record/ DLHE (HESA)
- NHS Patient Record
- Inter-Departmental Business Register (ONS)
- Hospital Episode Statistics (DoH)
- Individualised Learner Records (BIS)
- Electoral Register
- Local Authority records, e.g. of social care services
- Housing provider records (including housing associations), e.g. rent and tenant databases



---

# Administrative Data – Benefits

- covers the universe of individuals
- data collection unobtrusive
- offers substantial sample size
- no attrition with associated problems for inference
- information often is captured in a lot of detail (sometimes unknown to individuals)
  
- marginal additional cost for reuse
- already used for the production of official statistics



---

# Administrative Data – Benefits

- “Reliably” linked with survey data / other administrative data
- Up-to-date



---

# Administrative Data -- Problems

- no consent from data subject for research use
- legislative issues with sharing and linking
- multiple data controllers (multiple requirements for access) [decentralised statistical organisations]
- data “less prepared” than survey data for research [quality control]
- documentation for research purposes may be poor
- time sensitive
- coverage [incompleteness]
- policy changes => conceptual changes
- big
- costs incurred are not seen as costs avoided
- researcher can not control content

# Administrative Data Taskforce

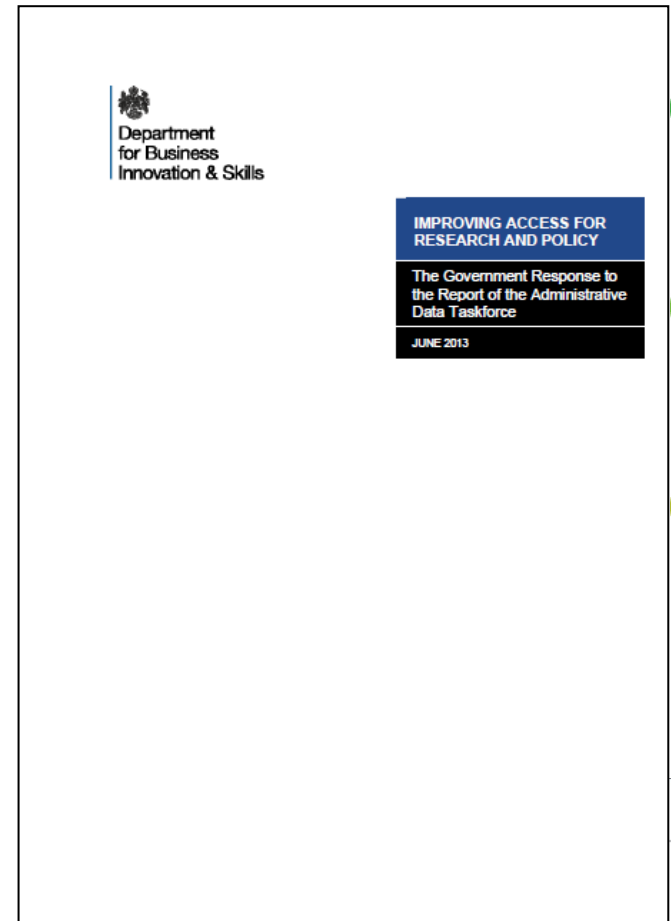
- Frontiers Meeting (May 2011)
- Administrative Data Taskforce (Dec 2011-Dec 2012)
- Administrative Data Taskforce Report (Dec 2012)
  - ADRC in each country of UK; an information gateway, a governing board
  - UK-wide researcher training and accreditation process
  - Strategy for engaging with the public
  - Generic legal gateway
  - Sufficient resources





# Government Response

“Data collected and held by Government is a unique resource. Unlocking that resource has the potential to develop new understanding and insights both between different fields of study and over time”



---

# Government Response

It emphasised the importance of:

- Building on existing activities, infrastructure and systems where feasible in developing a new UK-wide approach.
- Developing the infrastructure in a way that maximises the potential benefits to both government analysts and the wider research community. Both will ultimately benefit citizens.
- Ensuring that the full breadth of data sources held in administrative systems where they have analytical value are accessible for research purposes.

# Government Response (2)

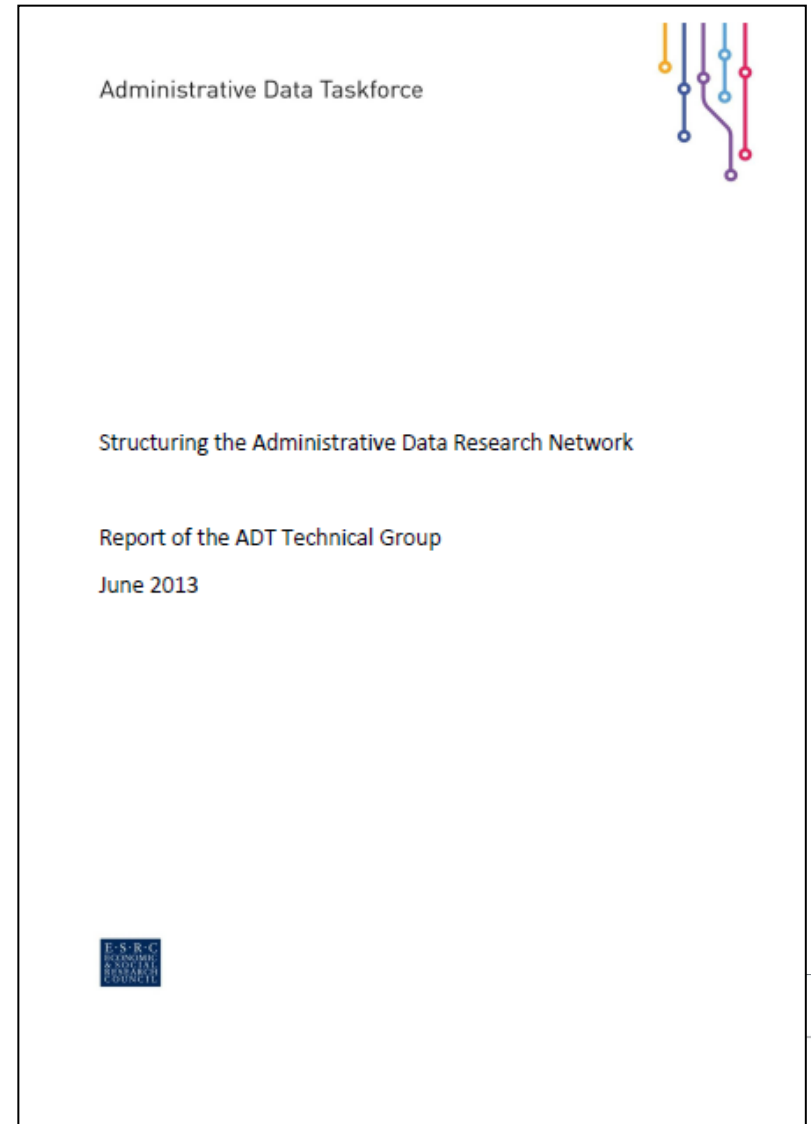
Also recognised that

- the devolved nature of some functions of government means that the infrastructure needs to reflect and accommodate the varying situations across the UK.
- data held by government varies in terms of its sensitivity,
- ADRCs could be asked to provide access to data another administration, thus development needs to maximise the value of investment, minimise the cost and burden and avoid duplication with existing facilities or other ADRCs.
- the ADRCs should extend and build on existing good practice, experience and expertise in the management of access to linked data, and ADRCs work very closely together.



# ADT Technical Group

- Structure of ADRCs
- Secure access procedures
- Researcher accreditation
- Research project accreditation
- Facilities accreditation
- Safe settings



# Government Response (3)

The screenshot shows the Economic and Social Research Council (ESRC) website. The header includes the ESRC logo and the tagline 'Shaping Society'. A search bar is located in the top right corner. Below the header, there are navigation tabs for 'Funding and guidance', 'Impacts and findings', 'Collaboration', and 'News, events and publications'. The 'News, events and publications' tab is active, and a sub-menu shows 'News', 'Events', 'Publications', 'Features and case studies', 'Videos', and 'Audio'. A breadcrumb trail reads: 'You are here: Home > News, events and publications > Press releases > The Big Data Family is born - David Willetts MP announces the ESRC Big Data Network'. The main article title is 'The Big Data Family is born - David Willetts MP announces the ESRC Big Data Network', dated '10 October 2013'. The text of the article describes the announcement of a £64 million investment in Big Data research centres and a data service, forming the Administrative Data Research Network (ADRN). A sidebar on the right contains a sign-up form for ESRC news, a 'My ESRC' section with 'Bookmark' and 'Print' buttons, and social media icons for Facebook, Twitter, and LinkedIn.

**Economic and Social Research Council**  
Shaping Society

Enter keywords to search ESRC

Funding and guidance | Impacts and findings | Collaboration | News, events and publications

News | Events | Publications | Features and case studies | Videos | Audio

You are here: ▶ Home ▶ News, events and publications ▶ Press releases ▶ The Big Data Family is born - David Willetts MP announces the ESRC Big Data Network

## The Big Data Family is born - David Willetts MP announces the ESRC Big Data Network

10 October 2013

David Willetts MP, Minister for Universities and Science, speaking at the prestigious Mountbatten Memorial Lecture last night, announced the recipients of the first phase of the Economic and Social Research Council's (ESRC) £64 million funding of investment in Big Data. Four new innovative administrative data research centres and a data service will strengthen the UK's competitive advantage in Big Data.

The centres and service together form the Administrative Data Research Network (ADRN) which will enable research based on linked data between government departments and be overseen by a single governance structure that will allow for consistent and robust decision-making.

At the core of the ESRC's Big Data Network are the centres to be led by the Universities of Southampton and Edinburgh, Swansea University and Queens University Belfast, with the administrative data service to be based at the University of Essex. Collectively, they will benefit from a grants package totalling approximately £34 million.

Sign up for ESRC  
Latest news from the  
Subscribe to our pu

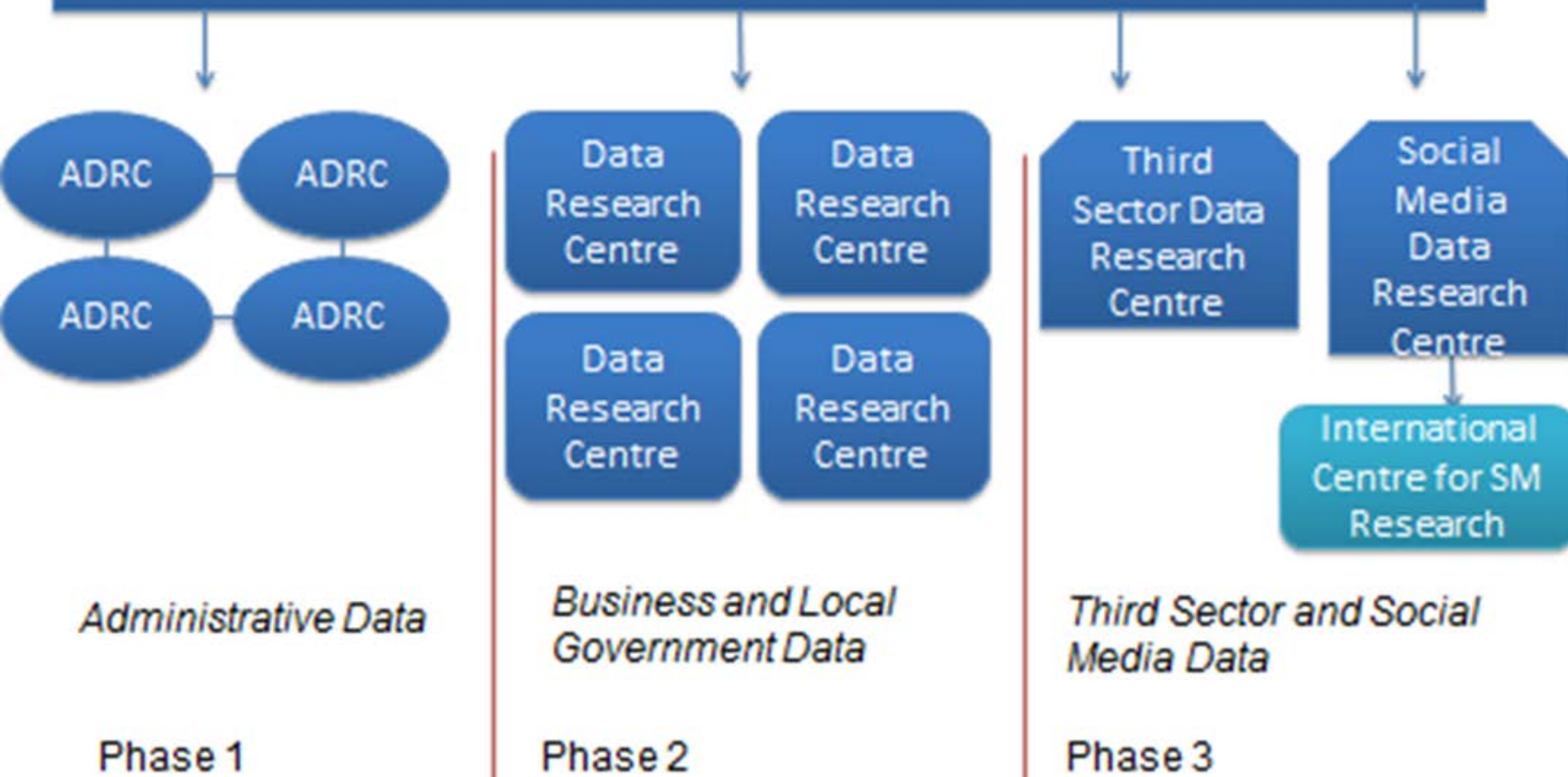
My ESRC  
This page:  
Bookmark Print

f t in

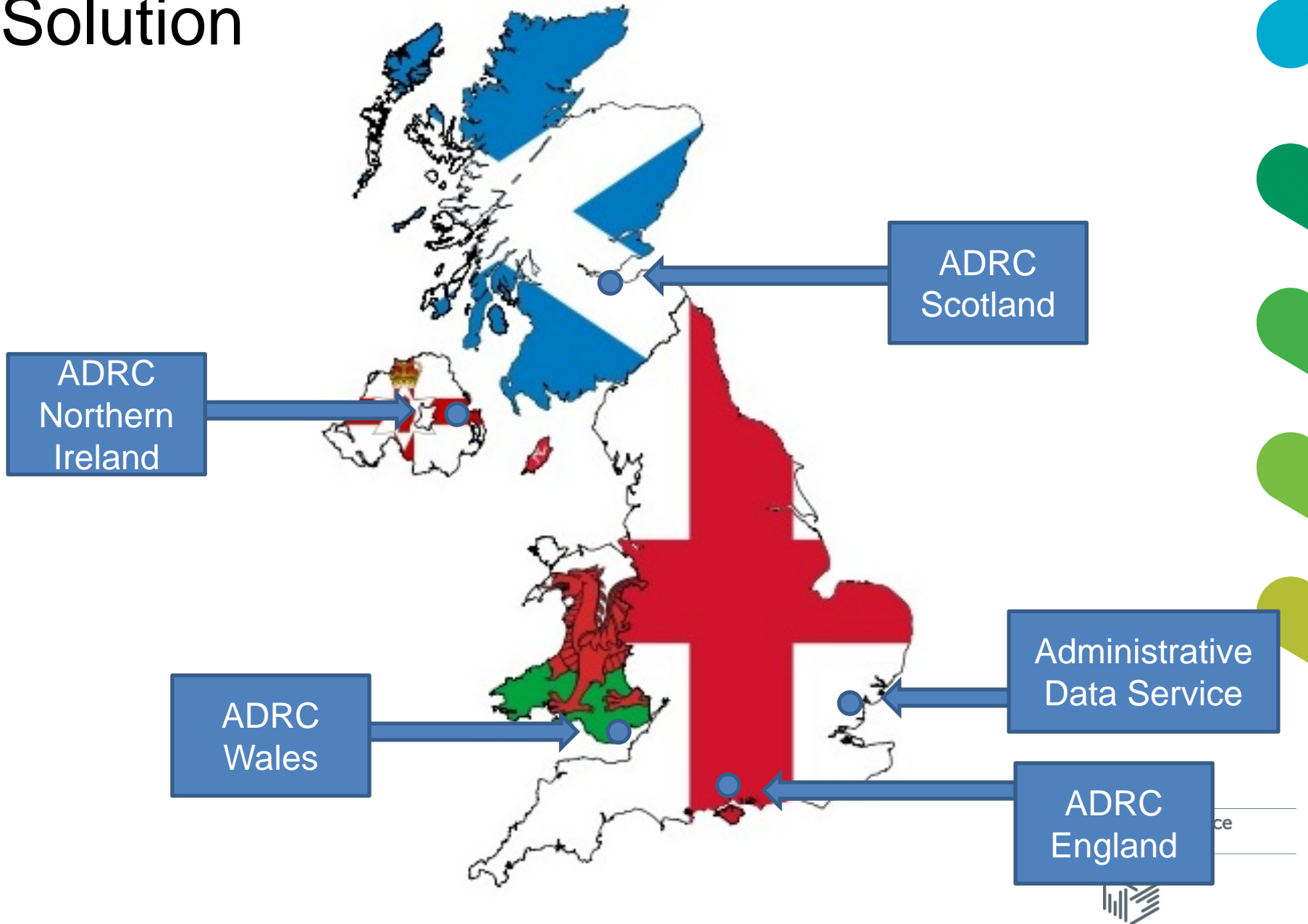
# ESRC Big Data Network

Governance Structures

UK Data Service and Administrative Data Service



# Solution



---

# Administrative Data Network

*To facilitate access to admin data held by government departments and enable linkage between the data.*

- An Administrative Data Service and Four Administrative Data Research Centres (ADRCs)
- Key partnerships with National Statistical Authorities
- A UK-wide governance for the network
- Safe settings
- Researcher training and accreditation
- Public engagement strategy



---

# Citizen participation

Citizens (data subjects) need to be aware:

- of how data about them is controlled;
- of how data about them is being used and for what purpose
- of how the results of these analyses are beneficial to society.

---

# Administrative Data Service

*To establish a seamless research infrastructure which places the UK at the forefront of research using administrative data resources*

## Key Activities

- Branding / Comms / Social Media
- Researcher engagement
- Public engagement
- Internal Coordination
- User service



---

# Key Activities

- Data and Linkage Commissioning
- Data Owner Engagement
- Trusted Third Parties
- Cross-sector standards' working groups
- Safe settings / Data confidentiality and SDC
- Metadata
- Researcher training
- Building student capacity
- Guides
- Legal issues



---

# Infrastructure issues

- National initiatives
- Networks vs Networked Infrastructures
- Limited provision for cross national data sharing
- Admin data is not driven by research issues per se
- Expensive information security requirements
- Methodologically underdeveloped [social science]
- Procedures unsophisticated
- Capacity building / data analysis skills
- Preservation / Destruction paradigms



---

# CONTACT

matthew@essex.ac.uk

Melanie Wright  
Director, Administrative Data Service  
melanie@essex.ac.uk

