

ESRC and the Administrative Data Research Network – background and context

Vanessa Cuthill 14 April 2014



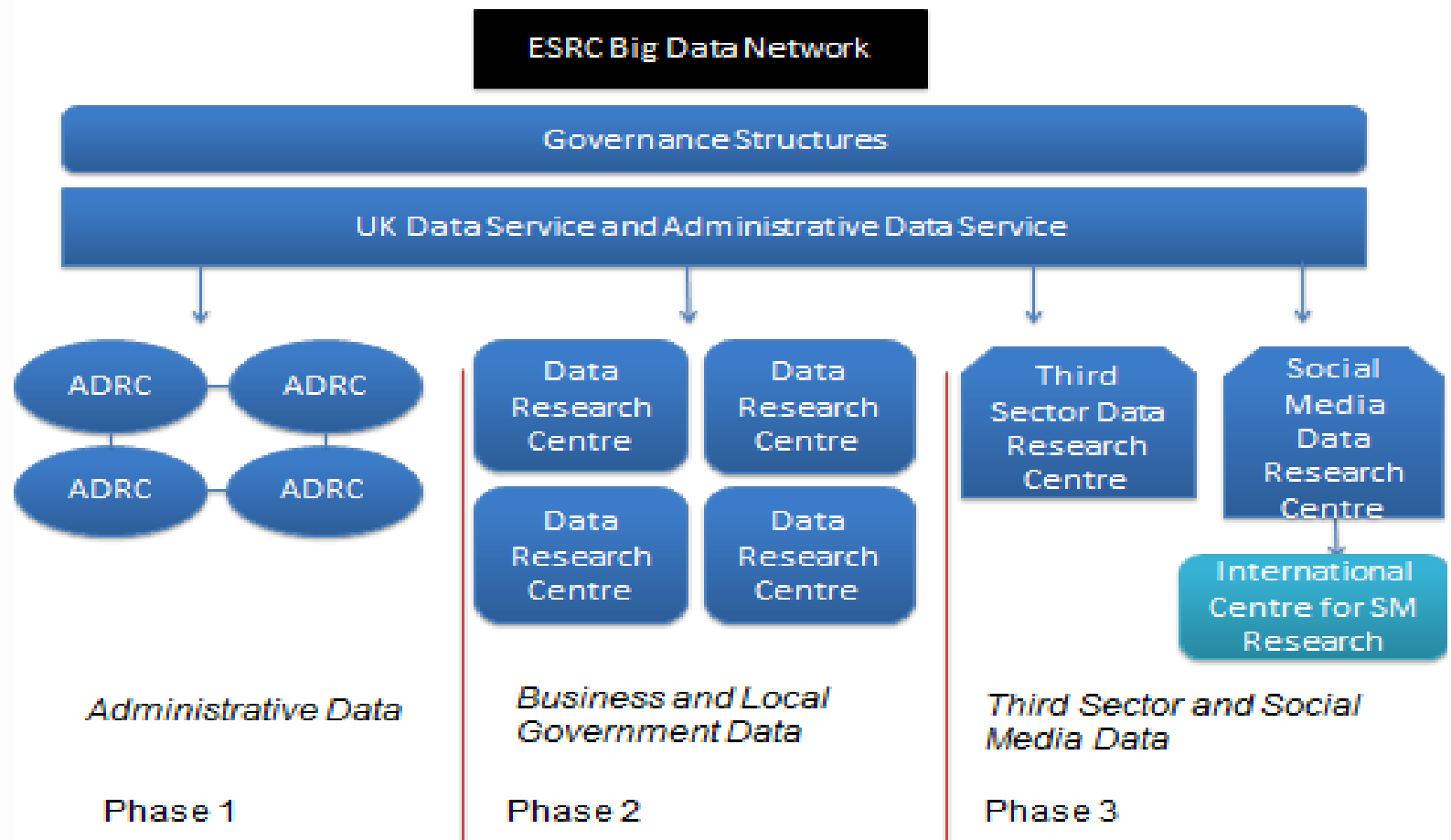
ESRC

- UK's major public sector funder of social science research and post graduate training (£200m budget)
- Non-Departmental Public Body, since 1965, largely funded through the Department of Business Innovation and Skills (BIS)
- Key Principles:
 - Quality
 - Impact
 - Independence

Capital Funding & Big Data

- Chancellor's Autumn Statement included £600m for science, research and innovation, £484m for RCUK
- Funds to support the development of innovative technologies across eight areas, including 'big data'
- Draws from RCUK Strategic Framework for Capital Investment (published Nov 2012)

Big Data Network



Background to the ADRN




“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 to the ADT Report 2013

ADRN – The Journey So far...

- 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 & accreditation process ***
 - ❖ **Strategy for engaging with the public ***
 - ❖ *Generic legal gateway*
 - ❖ *Sufficient resources*

The Journey Continues...

- 
- A decorative vertical bar on the left side of the slide, composed of horizontal segments of red, white, and blue.
- Chancellor's Autumn Statement (Dec 2012)
 - ESRC Technical Group (Feb - May 2013)
 - Government Response (June 2013)
 - ESRC Commissioning of ADRN (Jul-Oct 13)

Initial response from Government

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.

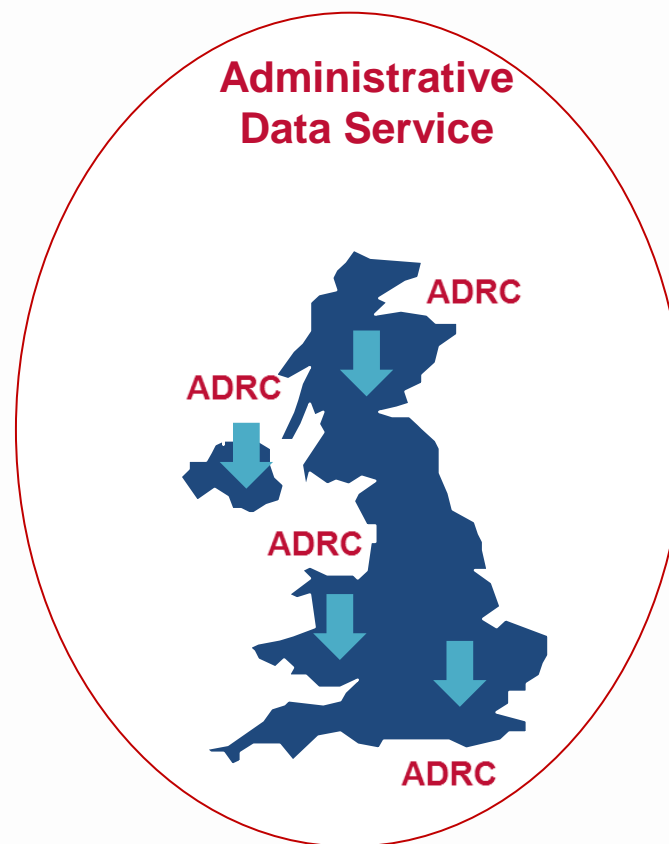
The ADRN

“ 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

ESRC investment ...

- **ADS** – Melanie Wright
Essex (UKDS / ADLS)
- **ADRC-E** Peter Smith
Southampton (Bloomsbury)
- **ADRC-NI** Dermot O'Reilly
QuB (Ulster)
- **ADRC-S** Chris Dibben
Edinburgh (Scottish
consortium)
- **ADRC-W** David Ford
Swansea (Cardiff)



Some benefits...

- Realising the opportunity to exploit the richness of administrative data for social and economic research and policy, and for better targeting of interventions.
- A step-change in procedures for access to, and linking between, such data across Government Departments
- Ensuring the UK remains at the forefront of research based on linked administrative data and is competitive on the international market
- Allowing cost effective re-use of existing data by avoiding the duplication of data collection and capitalising on past investments
- Enabling new methodologies and cutting-edge research
- Developing skills and capacity


Partnership

- Partnerships are key:

Across the ADRN universities;
National Statistical Authorities;
Existing infrastructure (UKDS, Farr, ADLS);
Data owners; and researchers...

“Realising this opportunity requires ... not just the development of a safe, secure and efficient system for linking, managing and analysing administrative data... but on further building of trust between data subjects, data providers, researchers and all other relevant parties” (ADT Report 2012)

Cross cutting areas

- 
- A decorative vertical bar on the left side of the slide, composed of horizontal segments in red, white, and blue.
- Branding and communications
 - ADRN operating model
 - Public engagement
 - Accreditation and training
 - Legislation
 - Governance

Key Components of the Governance

- ADRN Board

Independent assurance that the Network is delivering against core objectives (*strategic advice and assurance to ESRC; accountable to UKSA Board and to UK Parliaments*)

- Management Committee / Leadership Team

Advising the ESRC on the direction and management of the Network (*oversee the management of ADRN, including operational progress, finances, risks*)

- ADRC and ADS internal management

Operational management groups; advisory groups; ADRN-wide working and expert groups; Independent Review Panel

The ADRN Board

Rob Bumpstead

Deputy Director, Office of the Board and Chief Executive
UK Statistics Authority

The UK Statistics Authority

- Established on 1 April 2008 under the *Statistics and Registration Service Act 2007*, to promote and safeguard the production and publication of official statistics, for the public good.
- An independent body, operating at arm's length from government as a non-ministerial department, directly accountable to Parliament.

The UK Statistics Authority

The Authority's functions relate to its statutory areas of responsibility:

- Oversight of the UK official statistics system, and the promotion, safeguarding and monitoring of quality, comprehensiveness and good practice in relation to all official statistics;
- Production of a Code of Practice for Official Statistics and assessment of official statistics against the Code; and
- Governance of the Office for National Statistics, the UK's National Statistical Institute and largest producer of official statistics.

The Authority's strategic priorities

1. Quality
2. Impact
3. Efficiency
4. Coverage
5. Trustworthiness

The ADRN Board so far...

- A meeting of the Chief Executives of the UK Statistics Authority and the ESRC last summer.
- Continuing cooperation at working level.
- UK Statistics Authority Board agreed to accept the Taskforce's recommendations in October 2013.

Fit with the Authority's priorities

“Support the application of open data principles to official statistics.”

“The Authority will support the National Statistician, ONS and the wider statistical service in (i) achieving a step-change in the sharing of data across, and beyond, government for statistical purposes; and (ii) making data available in the most disaggregated form possible, and as soon as possible, for re-use by others.”

Establishing the ADRN Board

- In November 2013, the Authority announced publicly that it intended to accept the Taskforce's recommendation.
- The Chair was appointed by Authority Board.
- The Deputy Chair was appointed by the ESRC.
- Membership comprises a mix of nominated representatives and representatives appointed via open competition.

The ADRN Board's Objective

“The ADRN Board will promote and safeguard the linkage of administrative data for research and policy purposes that serve the public good.

It will provide assurance to Parliament and the public, through the Board of the UK Statistics Authority, about achievement of the Network's core aim: to facilitate linkage of routinely collected administrative data, thereby stimulating opportunities for innovative research and policy making.”

The ADRN Board's Role

“The ADRN Board will guide the strategic direction for the Network and will report and provide assurance to the Board of the UK Statistics Authority. The ADRN Board is responsible for providing independent assurance that the Network is delivering against its core objectives and is being developed, managed and maintained in a way that maximises its benefits.”

“It will also provide advice to the Economic and Social Research Council (ESRC) on these matters.”

Support to the Board

- Secretariat to the ADRN Board will be provided by the UK Statistics Authority.
- Dedicated ADRN secretariat resource has been newly recruited. Simon Whitworth will begin in the coming weeks.



Better Knowledge Better Society

Administrative Data
Service

Melanie Wright



Administrative Data
Research Network

An ESRC Data
Investment

Introduction

Partnership

- UK Data Service (Universities of Essex and Manchester) and
- Administrative Data Liaison Service (Universities of St Andrews/Edinburgh, Manchester and Oxford)

Five main teams:

- Data and User Services (Essex)
- Communications & Stakeholder Engagement (Essex)
- Training and Capacity Building (Manchester)
- Legal Guidance (Oxford)
- Management and coordination (Essex)



What We Do: Network Coordination

- Gluing together or herding cats?
- Coordinate activities across the Network
 - Facilitate internal working groups
 - Keeper of ADRN-wide Comms and PubEng strategies
 - Ensure synergy/efficiency not repetition
- Act as the front door to the Network
 - Website, branding
 - Help Desk, data catalogue, guides
 - Advice on legal issues
- Oversee/track the User Journey
 - Assist in project proposal preparation
 - User accreditation and training
 - Negotiate with data owners for access to data
 - Administer Independent Review Panel
 - Track impact
- Develop network-wide standards and protocols



Biggest Challenges

- Challenge to balance Service and Research
- Challenge to integrate across countries, sectors, established ways of working
- Challenge to manage public perception (especially in the wake of care.data)
- Challenge to find our place in wider data service landscape
- Challenge to assuage data owner anxiety





Better Knowledge Better Society

Administrative Data
Research Centre –
Wales

Professor David Ford



Administrative Data
Research Network

An ESRC Data
Investment

Introduction

- **Co apps** (Swansea, Cardiff, Bristol): Jones K (methods and IG), Davies (User Support), Jones IR (researcher liaison), Innes (criminology), Haines (Justice), Taylor (Education), Lyons (Public Health), Phillips (Older People), Blackaby (Economics), Mackie (housing), Pithouse (children)
- **SAIL Databank**: population data resource for Wales, (9billion recordings), inc. full privacy protection system & access mechanisms. TTP = NWIS
- **Welsh Government**: data providing and research agenda owners
- **Farr Institute@CIPHER**: also at Swansea, brand new shared data science building, deep collaboration, co-occupied with NHS and WG staff.

What we do

- **Privacy protection:** architectures, methods, tools, technologies, policies, procedures.
- **Data linkage**
- **Big data technologies**
- **Applied policy-relevant research**
- **Methodological advancement**
- **Supporting users**
- **Public involvement**
- **System-level solutions**
- **Project and programme management**

Key Challenges

- Pretending it's all easy and sorted
- Earning public trust (not care.data)
- Involving the public meaningfully
- Demonstrating researcher benefit
- Proving data safety
- Remembering data linkage is key
- Demonstrating UK reach
- Coordinating with Farr and other initiatives
- Remembering what we've got
- Staying imaginative and adaptable
- Getting down to the detail



Better Knowledge Better Society

Administrative Data
Research Centre –
England

Professor Peter
Smith



Administrative Data
Research Network

An ESRC Data
Investment

Introduction

- Consortium led by UoS with UCL, IoE, IFS, LSHTM, ONS
- Director: Smith (UoS); Deputy Directors: Gilbert (UCL), Martin (UoS)
- Management Group with representation from all institutions, all work-streams and senior IT staff: Dearden (IoE, IFS), Wilkinson (LSHTM), Denaxas (UCL), Durrant (UoS), Parchment (UoS), Humberstone (ONS)
- Senior Project Manager in Southampton; research, data support and administrative staff in: Southampton; Bloomsbury, London; ONS, Titchfield
- Scientific panel comprising all co-investigators: providing advice and leading coordinated research programmes



What we do

- Facilities including e-Infrastructure (Lead: Parchment)
- Data management and statistical support (Lead: Smith)
- Data resources (Lead: Martin)
- Research (Leads: Martin, Dearden, Wilkinson)
- Training, capacity building (Lead: Durrant)
- Working in collaboration with the ADRN (Lead: Smith)
- Public engagement (Lead: Gilbert)



Key Challenges & Opportunities

- To meet the expectation of users of the service, regarding range of data sources available and speed of access
- Legal and other barriers to data access
- To provide a unified as possible process for accessing linked administrative data across the ADRN
- Working with ONS
- To establish an internationally recognised centre for linked admin data research, part of a thriving UK network
- Early port of call in public debate about research using linked data





Better Knowledge Better Society

Administrative Data
Research Centre –
Northern Ireland

Dr Dermot O'Reilly



Administrative Data
Research Network

An ESRC Data
Investment

Introduction



What we do

- Linkage service
- Safe setting
- Data acquisition
- Public engagement
- Training, capacity building
- Research

Researcher support



Key Challenges & Opportunities

- Data custodians
- Public engagement
- Policy impact
- Smooth pan-UK operation
- Multi-disciplinary working
- Capacity building





Better Knowledge Better Society

Administrative Data
Research Centre –
Scotland

Professor Chris
Dibben



Administrative Data
Research Network

An ESRC Data
Investment

Introduction

- Brings together major Scottish centres: Aberdeen, Dundee, Edinburgh, Glasgow Herriot-Watt, St Andrews, NHS working closely with Scottish Government.
- Leading experts in the Law, linkage, computer science (eg Natural Language Processing, Machine Learning) and significant research areas
- Builds on existing services
- Significant programme of public engagement
- Will exploit Scotland's rich admin. data (historical and contemporary)



What we do

Provide:

- state-of-the-art facilities for research access to de-identified administrative data by accredited researchers
- data management and statistical analysis support functions for external researchers accessing the data

Develop the capacity to:

- Engage with the public and effectively communicate value/ worth
- Respond to the legal environment emerging out of cross-sectoral linkage
- Store, access and link large complex datasets exploiting new linkage methods
- Geocode administrative data from the 1850s
- Extract information from free text resources
- Produce synthetic versions of linked administrative datasets

Research:

- Incomes, benefits and labour market
- Informal care - the implications of caring for health and wellbeing
- Parental mental health and filial educational outcomes
- Lifecourses – by linking the Scottish Mental Surveys of 1932 (1921 Cohort) and 1947 (1936 Cohort)



Key Challenges & Opportunities

- Different constituencies
 - Very positive engagement with Scottish Government
 - Working closely with Farr Institute Scotland
 - ADRN network
- Building on existing resources – they are already there/ designed
- Public – confidence/ satisfaction
- Sustainability



Public dialogues on reuse of admin data for research

Melanie Knetsch – Deputy Head of Communications/PE lead



Background

- Throughout October - November 2013 public dialogues held in 7 locations across UK (Ipsos MORI led)
- The aims were to:
 - To better understand the cultural barriers around linking administrative data
 - To begin the process of creating a language that is meaningful and accessible to the public.
 - To test the public perceptions of the rules that ESRC ADRCs will be subject to and to provide the ADRCs with data on public attitudes and appetite for engagement
 - *(To provide ONS with more detailed evidence on public views of their current front-running option for Beyond 2011)*

Findings -Social research and data

- A lack of familiarity with the process and methods of social research = concerns about wasted effort.
- The key concern -whether social research actually leads to social improvement
- Personal data security very important - concerned about identity theft, and personal data being sold on to other organisations.
- Overall more trust in government's intentions than commercial companies on data security and protection from data misuse or fraud.
- Assumption that administrative data is already linked and shared across departments, = support for this for operational uses.

Findings - Administrative data

- Uncontroversial - use of administrative data to improve the way services are run or to increase national security
- Those **more trusting** of government tended to think that any government use of data would be fairly benign and benefit the general public, or a subsection of the population.
- Strong support for the use of administrative data in planning for future service provision.
- View that data should be used to reduce fraud in government services.
- Open to data being used more efficiently to avoid repeat collection of the same information from individuals.

Administrative data - concerns

- Too much unnecessary personal data collected and duplicated by departments and services.
- It is not always clear why some information is collected when it does not seem directly relevant to the service received.
- The potential for profiling certain subsections of the population, or individuals, and potentially treating them differently on that basis.

“Sometimes I think [the government] collect all your details because they want to be pigeonholed. Like in the computer world they can put your name in and see what your likes and dislikes are.” King’s Lynn

Admin data - concerns

- Administrative data could be inaccurate, especially where it is self-reported = negative consequences for individuals or groups.
- Using linked administrative data to justify and implement controversial policies, such as the bedroom tax:
“They found out a member of my family had passed away because of the electoral roll, so now I pay bedroom tax.” King’s Lynn
- Government departments holding prejudicial administrative data (e.g. records of arrests) for longer than they should, and sharing it that is data that is **taken** vs that which is **given**
- The factors that influence what data is collected by government, which in turn influences the way data is interpreted and the decisions government makes. (eg data about schools and pupils - the consequences of data-driven targets)

The ADRN initiative

Concerns	Reassurance points	Remaining questions
<ul style="list-style-type: none"> ❖ De-identification – will it be possible to identify people even without their ‘personal details’ ❖ De-identification process – is this secure? Where does the personal data ‘go’? ❖ Security – is it possible to hack? Does that even matter? ❖ Cost – is this £34 million well spent? 	<ul style="list-style-type: none"> ✓ ADRCs’ independence ✓ Data could improve public services or allocate funding better ✓ Efficiency – Save money, or reduce burden on population ✓ Better regulation than current system ✓ No ‘super-database’ ✓ Limitations on researchers 	<ul style="list-style-type: none"> ? Can this improve operational use of administrative data? ? How much have government departments ‘bought in’ to this idea? ? Who works for the ADS and how are they governed? ? What will be the punishment for any misuse of data by researchers?

Support for initiative IF:

1. The data is linked for socially beneficial purposes
“As long as it’s used for good, like to develop things, improve services, improve knowledge.” Belfast
2. It is fully de-identified – partial vs full postcodes
3. It is kept secure at all times – concerns around remote access
4. No commercial gain for business including commercial access

BUT participants needed extensive information and discussions with experts and researchers in order to be satisfied that these conditions would be met under the ADRN plans.

So - simply publicising these three conditions may not be enough to ensure that the general public are reassured about or support the work of the ADRN.

Admin data – communication challenges

People's views change with new information

- ***“Speaking to [ADS representative] has really helped, someone within the service who is passionate about it.” King’s Lynn***
- ***“Now I know more, and it’s de-identified, a lot of the bad factors have been cleared up like trust issues.” Cardiff***

Admin data – communication challenges

- Showing the value = IMPACT

Strong view that those applying for data to be linked means that “something has to come out of it”

(improving understanding of society or services)

- Security of data = researchers who use and remote access

“Now I know more about the researchers’ permissions being removed they’ve got more motive to keep safe, it’s personal to them.” King’s Lynn

Admin data – communication challenges

- Keeping and storing data - participants tended towards one of two views:
 - 1) Linked datasets should be stored so that researchers - could have access after the initial research project is complete.
 - 2) Destruction of data after project completion reassuring

Next steps

- Report shared with ADRC/ADS
- Looking at ways to share report
- We are developing two short animated videos to help explain:
 - a) What is social science research
 - b) Introduction to data linkage