



**Department
for Business
Innovation & Skills**

**IMPROVING ACCESS FOR
RESEARCH AND POLICY**

**The Government Response to
the Report of the Administrative
Data Taskforce**

JUNE 2013

Contents

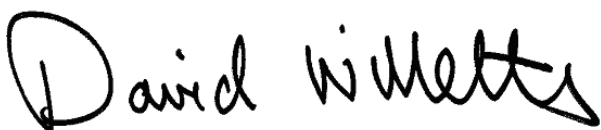
Contents.....	2
Ministerial foreword.....	3
Introduction	4
Government Response to Individual Recommendations	6
R1 An Administrative Data Research Centre (ADRC) should be established in each of the four countries in the UK.....	6
R2 Legislation should be enacted to facilitate research access to administrative data and to allow data linkage between departments to take place more efficiently	12
R3 A single UK-wide researcher accreditation process, built on best national and international practice, should be established.....	13
R4 A strategy for engaging with the public should be instituted.....	14
R5 Sufficient funds should be put in place to support improved research access to and linkage between administrative data	15

Ministerial foreword

We welcome the Administrative Data Taskforce's report as an important contribution to positioning the UK at the forefront of the use and reuse of de-identified administrative data for research and policy purposes. The UK Government would like to thank Sir Alan Langlands and Professor Paul Boyle for the work they have done in bringing together funders and others to consider this important issue.

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. Advances in analytics have improved capacity to analyse large datasets and to open up new seams of inquiry, improving the evidence base and leading to better policies, better outcomes and potentially economic growth. Over recent years, more data held by government has started to be released and the report points to many examples where new insights have come from existing data analysed in new ways. The ADT set out some concrete actions to speed up the process by which the full potential of administrative data can be exploited for research and statistical purposes.

The Government agrees with the overall direction of travel indicated in the ADT report and supports the broad recommendations and ambitions. We are committed to working constructively with the research community in developing proposals for taking them forward. This document is a first response following initial consideration within Government of the proposals. We are already working with the ESRC, the Devolved Administrations and others in developing more detailed plans.

A handwritten signature in black ink that reads "David Willetts". The signature is written in a cursive, slightly slanted style.

David Willetts MP

Introduction

The Administrative Data Taskforce report *Improving Access for Research and Policy* was published in December 2012. The Administrative Data Taskforce (ADT) was formed in December 2011 and was chaired by Sir Alan Langlands. The work of the group was led by Professor Paul Boyle, Chief Executive of the Economic and Social Research Council (ESRC), drawing on the expertise and views of a range of experts.

Both the desire and the capacity to exploit large data collections for research and policy purposes are growing in tandem. The *Open Data White Paper: Unleashing the Potential* (June 2012) set out the ambition to ensure effective use and reuse of data. More recently, *The Shakespeare Review An independent Review of Public Sector Information* (May 2013) <https://www.gov.uk/government/publications/shakespeare-review-of-public-sector-information> looks more broadly at the use and reuse of public sector information and the Caldicott2 Review *Information: to share or not to share? The Information Governance Review* (April 2013) <https://www.gov.uk/government/publications/the-information-governance-review> has recently reported on Information Governance within the Health Service. All support the appropriate use of administrative data held by government to facilitate research and policy development.

The ADT looked specifically at facilitating the reuse of administrative data for research and statistical purposes by bona fide researchers. Dame Fiona Caldicott's review made recommendations on the appropriate balance between the protection of confidential information within our health and care records and the use and sharing of information to improve the quality and safety of our own care and for the benefit of wider society. The Government will respond to the Caldicott 2 review later in the summer. That response will, in turn, have a bearing on how the Government takes forward the recommendations of the ADT, in particular, in relation to personal health and care data, and the need to respect the privacy and wishes of the patient about some of the most sensitive information held about them.

Government collects and holds a vast amount of data as part of the normal transaction of government business. Similarly, government collects data for the purpose of producing statistics about the current state of the economy and society. The ability to link and analyse data held by government has the potential to add new insights to our understanding of how society and the economy performs and to reduce the need for separate data collections where we ask, for statistical purposes, for the same information that has already been provided for administrative purposes. This would be of value both within government and to the wider research community and therefore ultimately to the citizen.

The ADT report makes a number of recommendations which have the potential to deliver a significant step change in harnessing the full potential of administrative data to illuminate policy options and to monitor progress in policy delivery. There already exist a number of initiatives which have demonstrated the potential of linked data to inform policy and social debate. Many of these are cited in the report and demonstrate the potential of administrative data to add new insights to the knowledge base. For example, linking

administrative data on students has provided new information on progression through the various stages of the education system and potentially into the labour market. Similarly, linked data has contributed new understanding in the fields of crime and offending, in health and life expectancy and disease. Several government data owners have already set up facilities to allow greater use of administrative and survey data for research and statistical purposes through the establishment of data laboratories or legal gateways. While a need for link-specific gateways will continue to exist, the ADT proposes that we pull this effort together and take it to the next level.

In general terms, this initial response from Government emphasises 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 to Individual Recommendations

R1 An Administrative Data Research Centre (ADRC) should be established in each of the four countries in the UK

R1.1 The ADRCs will be responsible for commissioning and undertaking linkage of data from different government departments and making the linked data available for analysis, thereby creating new resources for a growing research agenda. Analyses of within sector data (e.g. linking medical records between primary and secondary care) and linking of data between departments for operational purposes may continue to be conducted by the relevant government departments and agencies.

R1.2 Each ADRC must be a state-of-the-art facility for research access to de-identified administrative data. While the detailed organisation and structure of each centre may reflect national variations in access arrangements, the fundamental common feature of each ADRC will consist of a secure access facility that meets the most stringent international standards. These should be attractive research environments (i.e. powerful hardware and analytical software should be available together with access to relevant metadata supplied by the data owners). Access to data and methodological and statistical support should be free to bona fide researchers who have gained accreditation status (see recommendations 3.1 to 3.3).

R1.6 Access to data held in the ADRCs by accredited researchers will be possible using three approaches. For all of these, no individual-level records will be released from the ADRCs. First, researchers can visit the ADRC secure data access facility, where their analyses of the relevant data sub-set will be overseen by the ADRC support team. Second, researchers can submit statistical syntax to the ADRC support team who will run the analysis on the dataset on behalf of the researcher (results would be thoroughly checked before return). Third, remote secure data access facilities may be established which allow virtual access to datasets held in the ADRCs. With the latter approach, no data would be transferred to these remote safe settings, which would use state-of-the-art technologies and apply rigorous international standards, equivalent to those used in the ADRCs themselves, to provide a secure environment for researchers to undertake their analyses.

The Government agrees that appropriate infrastructure needs to be in place to ensure safe and trusted access for the purpose of analysing de-identified or pseudo-anonymised linked microdata. These data provide the necessary granularity for the types of research purposes described in the report as opposed to aggregate data. They do however have a re-identification risk which is why having the suggested protocols and security arrangements in place are essential.

The devolved nature of some functions of government means that the infrastructure put in place will need to reflect and accommodate the varying situations across the different parts of the UK. Equally, data held by government varies in terms of its sensitivity and the infrastructure will need to accommodate data with different Business Impact Level ratings appropriately and securely. As any one of the ADRCs could be asked to provide access to data from another administration, the development of the ADRCs needs to be done in a way that maximises the value of investment, minimises the cost and burden and avoids duplication with either existing facilities or with other ADRCs. To achieve this, the ADRCs should seek to extend and build on existing good practice, experience and expertise in the management of access to linked data, of which there is a considerable amount, and ADRCs will inevitably need to work very closely together.

The Government welcomes the emphasis on operating to the highest international standards. This initiative offers an opportunity for the UK to take the lead. The safeguarding of data is of deepest importance as is the assurance that the research is conducted in a secure environment and the outputs from such research are not disclosive in any way.

To maximise the utility and benefit of making linked data available for analysis, it will be important that access to the data is as flexible as possible within the safeguards required. We support the approach of having multiple means of accessing the resource so that physical location does not become a barrier and that no disadvantage accrues to some potential users simply due to location. All means of accessing the data must be robust in ensuring the security of the data and the outputs. There are already a number of models of existing good practice on which to build such as those established by the Virtual Microdata Laboratory at the Office for National Statistics (ONS), the data laboratories of other government departments, the ESRC funded Secure Data Service and UK Data Service, and in Scotland, Wales and Northern Ireland respectively the Scottish Informatics Programme (SHIP) and the Scottish Longitudinal Survey (SLS), the Secure Anonymised Data Linkage (SAIL) databank and the Northern Ireland Longitudinal Study (NILS) resources. For Health and Social Care there are already arrangements in place to provide access to data. These have been further strengthened from April 2013 when the Health and Social Care Information Centre (HSCIC) gained increased powers to provide data linkage services that can be used by others.

The ADRCs will provide an opportunity not just for legitimate and approved researchers to access the data but will also be attractive to analysts within government. From the government user perspective, the needs and benefits of using administrative data for research and statistical purposes are broader than simply linking data on individuals. There is equal interest and value in making available linked de-identified data on other units of analysis such as businesses, employers or on entities such as properties within a secure environment. For example our understanding of pathways from education into employment would be further enhanced, in addition to the linkages given in Box 2 of the report, if linkages were also made between employees and employers to better understand job creation, wage dynamics and economic growth.

The value of including data from outside of government is also demonstrated in the report (Box 4) where the value of data held by energy companies has been used to support activities such as the decennial census. Further uses for data from such sources include

the ability to be able to model energy usage by linking property characteristics with energy consumption and household characteristics. This type of valuable analysis could all be made easier through a structure such as an ADRC.

The current legal restrictions on sharing data for research and statistical purposes, as well as the considerable practical difficulties, even within government, act as a constraint on what research or analysis is possible on important issues of the day. In some areas, such as health and social care, there are existing powers for linking, sharing, and publishing data in safe, anonymised formats delivering research beneficial to improving and safeguarding public health. However in other fields it is more difficult. Notwithstanding that providing access to de-identified linked individual data will deliver immediate benefits in broadening analytical scope, the aspirations for the ongoing development should be to broaden the reach of the ADRCs to include data on different entities where appropriate and relevant. Sufficient similarities exist in the techniques and methods employed, as well as the underlying analytical capability, to suggest that a single approach could be developed that accommodates all types of administrative data regardless of the unit of observation and with whatever other data is relevant and available, including data from some private providers.

R1.3 Personal identifiers (names, addresses, precise date of birth, national insurance numbers, etc.) attached to administrative data records will not be available to, or held in, the ADRCs; hence, both ADRC staff and researchers accessing data through ADRCs will not have sight of such personal identifying information. Linkage will be achieved through the use of third parties who have the expertise to provide secure data linkage services for matching personal records from existing data systems.

R1.4 The ADRCs must maintain a full audit trail of all activity relating to data access and linking. This should include the establishment of monitoring and performance indicators. They must operate output control systems that provide for the vetting of all research outputs for confidentiality issues.

The Government supports the proposed approach of using third party indexers and the reassurance that personal identifiers will not be available to, or held in, the ADRC. This separation of function is welcome as it will maximise the protection of privacy and minimise the risk of any inadvertent disclosure as only the data required for linking will be made available to develop the index and no identifying information will be held alongside the variables for analysis.

We also welcome the strong emphasis on maintaining a full audit trail. The confidence of data owners and data subjects will be important in bringing the full benefits of linked data to fruition. All the necessary protocols and safeguards must be in place, operated to the highest standards and fully documented.

R1.5 The ADRCs must have the capacity to conduct original research using these data, as well as exploring issues relating to data linkage methods, the quality of linked data, its coverage of specific populations and its suitability for particular research purposes. In addition to their own research function, the ADRCs will engage in training and capacity-building, supporting researchers who gain permission to undertake research on linked administrative data. By these means the ADRCs will contribute to the development of good practice for research access to de-identified personal data.

Developing capability and increasing capacity to undertake robust and insightful analysis of linked data will be important for maximising the potential of data for statistical, research and policy purposes. In particular, developing appropriate measures of data quality alongside the development of methods for linking and analysis will be important, particularly as this work gains currency. As this is an emerging area there will be significant need and opportunities for new methods and techniques to be developed to support the use of administrative data for research and statistical purposes. We believe the ADRCs will be best placed to deliver the relevant methodology and development work involving appropriate experts and would wish to see that as a major focus of their programme of work in the initial stages.

The Taskforce recommendations relate to the opportunities presented by analysing existing data. However, insights will emerge during analysis around how the associated data collection could be improved, either through standardisation of classifications or improvements in coherence between sources, or the identification of powerful additional explanatory variables for which additional data could be collected. As expertise and experience is gained in analysing linked administrative data, where feasible and appropriate, advantage should be taken of opportunities to improve the data quality making it a richer and more useful resource going forward.

Organisations appointed as ADRCs will no doubt adopt different administrative structures. It will be important that an institution involved in or hosting an ADRC does not accrue any competitive advantage when bidding for other research contracts that involve analysing data accessed via the ADRC or provide privileged access to linked data. It will be for the Governing Board to ensure that a level playing field is in operation such that all access for externally funded research is approved through the same application and approval procedures as any other researcher.

R1.7 A UK Governing Board will be established to provide the governance structure for the ADRCs. Together with the directors of the ADRCs, representatives from the funders and international experts in the field of administrative data use, membership of the Governing Board should include senior representatives from the relevant government departments, agencies and devolved administrative bodies that provide access to their data through one or more ADRCs. At least one lay member will also be appointed.

R1.8 The Governing Board will perform two key functions. At the strategic level it will have a focus on leadership and enablement in order to promote and facilitate safe research on administrative data for public and policy benefit. It will commission work and co-opt members if necessary to provide expertise in areas relating to ethical standards, international experience, linkage methodology (including linking administrative data to existing surveys and longitudinal studies), safe setting security, legal and other relevant issues.

In terms of its day-to-day work, a sub-group of the Governing Board will liaise closely with government departments and agencies to approve requests for access to and linkage between administrative datasets, as well as linkage to other non-administrative datasets. The Governing Board will also work to encourage collaboration across the ADRCs, and with other research data centres. To assist with this collaborative approach and to

encourage cross-national working, the ADRCs will report on their activities on an annual basis to the Governing Board.

R1.9 The Governing Board will report on an annual basis to a body responsible to the UK Parliament which will monitor progress. We recommend that the UK Statistics Authority or another similar body fulfils this role. The report will include a set of relevant performance indicators and will detail achievements; obstacles encountered and proposed developments for the following year.

Given the nature and sensitivity of accessing de-identified linked data and the need for robust stewardship of the activity, it will be important to have authoritative and appropriate governance arrangements in place. We are supportive of the proposal that the governance should be broad ranging, including representatives of the various stakeholders as well as the lay community. Equally, we are supportive of the proposal for the Governing Board to ensure co-ordination across the network of ADRCs and to report annually on their activities.

The final shape and composition of the Governing Board will need to be developed in light of the final or emerging shape of the ADRCs themselves. The Board should be sufficient in size to ensure adequate and appropriate oversight and scrutiny of the developments while at the same time minimising burden on members. We anticipate that the arrangements will likely need to evolve over time as the network matures and more experience is gained in the management of access to linked data. The Governing Board arrangements must also ensure that any potential conflict of interest is avoided with the necessary checks and balances in place, for example to ensure appropriate processes for separation of function where data owners are also data users. Nonetheless, the responsibility for allowing access remains with the departmental owners of the data.

The Government supports the principle of the Governing Board reporting to Parliament via an independent body. With independence there will be constraints on membership of the Governing Board which may be incompatible with the desire to ensure representation of data owners and data users within government. Given the broad reach of the activity and the high relevance of the proposals for the statistical and research communities within government, we agree that the UK Statistics Authority could be well placed to provide appropriate oversight. As the final shape of the ADRCs is still in development the UK Statistics Authority Board will wish to consider the appropriateness and impact of taking on these new responsibilities. The reporting structure will need to be suitable for the eventual format and organisational composition of the ADRCs which has yet to be agreed.

The UK Statistics Authority will make its own response to the ADT report.

R1.10 An important part of the work of the ADRCs will be to ensure that proposals for access to and linkage between administrative datasets are legal, viable, and technically feasible and that the research skills necessary to conduct the research efficiently are available. Information about the ADRCs, including access arrangements, will be managed through an Information Gateway. This single web portal will be used to manage applications to use data and will provide detailed information about the administrative data that are available, the metadata attached to each of these datasets, and the results and impact of studies conducted through the ADRCs. The Information Gateway will likely be managed by one of the ADRCs.

The Government supports the idea of a single gateway for access to information about administrative data across the UK and a single portal for access to linked data. The gateway should aim to link in with other initiatives that help users understand what data is available and the different arrangements through which access to data is provided. Centralising this function will have a number of benefits. It will reduce the burden on researchers in identifying data sources but will also enable co-ordination of activities across the full range of topics and administrative boundaries and will assist in monitoring take-up. Information about take-up will also inform the work to ensure that the necessary capacity and capability are developed across the analytical community.

The Government is keen to ensure that there is no unnecessary duplication of effort either between data owning organisations and the ADRCs or between the ADRCs themselves. The burden on data owners should also be kept to a minimum. As legal responsibility for sharing its data rests with the data owning organisation these checks will continue to fall to the relevant departments.

R1.11 The ADRCs will provide access, free of charge, to government administrative data by publically-funded researchers, including those funded by or working on behalf of charities and the third/voluntary sector. The ADRCs will not handle commercial data, or consider private sector requests for access to and linkage between administrative datasets held by public sector bodies. However, the Taskforce recognises that there could well be potential benefits that derive from private sector data and related research interests. The Governing Board will, at an early stage, investigate guidelines for access and linkage by private sector interests, as well as commissioning public engagement work on this topic.

It will evaluate the public benefits that will derive from specific requests for access, the need to ensure transparency in the research process, balancing these against the potential risk of reputational damage to data holders that might arise from the public perception of specific requests for access. In so doing, the Governing Board must note that there will be no derogation of the authority that government departments have in deciding how their data might be used in specific projects.

R1.12 An external review should be conducted after these arrangements have become embedded. This review would examine conditions of access for different groups of research interests, both public and private, with the aim of ensuring that these have evolved in a fair and open manner and with due regard for the views of the public.

The Government supports the proposition that the ADRCs should provide access to linked de-identified data for publically funded researchers and those from the third/voluntary sector. Of equal importance is for the ADRCs to meet the need of government analysts to monitor progress and support policy development.

We agree with the recommendation that private sector requests for access to the ADRC should be out of scope at this initial stage of development. However, there are clear synergies in methods and technologies with developments in Open Data and Big Data that would mean, in time, that interests are likely to converge. In developing the ADRCs it would therefore be useful to bear in mind the potential future landscape in what is a fast-moving environment so that private sector involvement could be facilitated in the longer

term if considered appropriate (either as a provider of data or as possible accredited research user).

The report made reference to the value of private sector data such as that held by energy companies (Box 4) and where there is a need to include commercial or private sector data for linking for clear analytical purposes, then this should be decided on a case-by-case basis. There are existing examples from both the health and energy arena where linking with data from the private sector is already the practice and lessons can be learned from experience elsewhere. There are other initiatives which consider the value of public sector information which are outside the scope of the ADT. There will be many scenarios to be worked through in practice and the limits initially will be driven as much by practical considerations, including resources and capacity, as anything else. The ambitions for the scope should be broad in the longer term and we support the recommendation that the Governing Board will investigate access and linkage by the private sector.

We support the recommendation to review the arrangements once they are up and running. New activities such as this, while building on considerable existing experience and expertise will necessarily encounter new issues due to scale. As the facilities grow in breadth and experience it will be essential to test how the procedures work for all parties concerned and to demonstrate where the benefits are realised as a result of analysing linked data. Additionally the value of publishing or making available aggregate statistics generated from linked administrative data will need to be considered as the activity develops.

R2 Legislation should be enacted to facilitate research access to administrative data and to allow data linkage between departments to take place more efficiently

R2.1 Government departments, Local Authorities, agencies and other public bodies face different legal restrictions on the nature of the access they might provide for research using data they control. In such cases, specific legal gateways have been established to resolve this issue.

However, recent experience demonstrates that current link-specific gateway legislation is both cumbersome and inefficient. The Taskforce recommends that primary legislation should be sought to provide a generic legal gateway for research and statistical purposes that enables efficient access to, and linkage between, administrative data held by different government departments, agencies and other statutory bodies.

R2.2 An agreed set of ethical standards should be produced, drawing on well-established ethical guidelines and covering the research uses to which administrative data (and administrative data linked to other types of data, including surveys) may and may not be put.

R2.3 In situations where linkage is proposed between large and de-identified datasets, consent for linkage is not required under the Data Protection Act 1998. Where linkage involves the addition of administrative data to information collected by survey methods, it is

both practicable and desirable to request consent for linkage from data subjects, even though the linked data will be de-identified prior to analysis. A common approach to the method of obtaining consent will be developed which will improve the efficiency of consent procedures and permit wider sharing of such linked data for research purposes.

The rules governing what data can be shared and for what purpose vary across data holdings and across administrations. Some data can be shared under common powers, some via link specific gateways and other data is governed by specific legislation. The absence of a single set of rules adds to the complexity of gaining agreement to share and link data for research and statistical purposes. As a first step it would be helpful to map out where existing powers are sufficient and where there are specific barriers to address so as to inform the need and scope for further legislation. Simplification and standardisation of the legal framework would make it clearer as to what is and isn't allowed to be shared or linked. The constraints faced with third party intellectual property rights will need to be considered along with any Database rights. It will also need to be aligned with developments on public sector information such as outlined in The Shakespeare Review (May 2013) and are subject to the Re-use of Public Sector Information Regulations 2005. The benefits that will accrue from making better use of administrative data for research and statistics will be many and will be across the whole of the policy spectrum as evidenced in the report. Some of these benefits can be delivered under existing powers. We will consider what legislation might be required to make access to de-identified administrative data for statistics and research purposes available in an appropriate and secure setting.

R3 A single UK-wide researcher accreditation process, built on best national and international practice, should be established

R3.1 All researchers wishing to analyse administrative data through the ADRCs should be required to gain accreditation status.

R3.2 A single accreditation process needs to be designed which builds on best national and international practice and is acceptable to all UK holders of administrative data.

R3.3 This accreditation process will require short course training including: best practice methods for working in a secure data environment; legal and ethical issues related to the use of individual records; methods of disclosure control; and an introduction to the variety of types of routinely collected administrative data.

R3.4 To retain accreditation status, researchers should be required to undertake an online update course on a regular basis, possibly annually.

R3.5 Where there is no recourse to legal penalties for any breach of data access conditions, accreditation will be accompanied by a set of sanctions to be applied in a proportionate manner to researchers and/or their institutions in the event of any such breach of the conditions of access to administrative data.

The Government supports this recommendation for access to be controlled through a process of approval and for this to be standardised. In line with current practice we consider it important that both the research project and the researcher are approved. This builds on existing good practice and will ensure that the same standards are applied across the full system.

Aside from academic researchers, an important group of potential users will be analysts working within government. While existing agreements and gateways exist for sharing some data for statistical and research purposes within government, these are often limited in scope and time consuming to put in place. Aside from the potential to incorporate more and different linked data, the ADRCs' work on methods and metadata development will make ADRCs an attractive proposition for government analysts, particularly in exploring and developing new analytical fields of inquiry. Existing and established programmes of analytical work involving linked data, where an existing gateway and facilities are in place, will likely continue in their current arrangement at least initially.

Building capability and expertise will be important pillars to support this work as will ensuring that researchers are regularly refreshed on their responsibilities in accessing such data. A strong, centralised researcher approval process will be essential to satisfy data owners that the systems put in place are sufficient to protect the reuse of their data.

R4 A strategy for engaging with the public should be instituted

R4.1 The Taskforce recognises the need to ensure that members of the public are aware of these developments in access to administrative data for research purposes, and that public concerns are given due attention. A strategy should be developed which will provide readily accessible information about the benefits of improved access to and linking between administrative data, and the measures being enacted to minimise risks of disclosure and to prevent inappropriate use of such data. The strategy should encompass procedures for raising public awareness about the need for research based on administrative data, the nature of public consultation and the roles that members of the public could play in decision-making regarding the administrative data to be accessed and linked.

R4.2 To help develop this strategy, the Governing Board, which will include lay members, should steer the plans for public engagement prepared by each of the Administrative Data Research Centres (see recommendation 4.4), ensuring that there is no unnecessary duplication of efforts and suggesting ways in which these activities associated with these plans may be undertaken in the most efficient and effective manner.

R4.3 To assist with its implementation the Taskforce recommends that each of the proposed ADRCs should appoint a public engagement and communications officer, to lead engagement with a wide audience and promoting dialogue about the research benefits that could accrue, the safeguards that are required to prevent any misuse of data and how both could be effectively and collaboratively achieved.

R4.4 The ADRCs should produce plans for public engagement. They should collaborate to plan and hold public events to explain the work they are doing and to generate debate

about the academic and broader social and economic benefits that derive from research using administrative data and the measures taken to ensure that the identities of individuals cannot be revealed.

The public acceptability of releasing linked de-identified data for research and statistical purposes in a controlled environment will be central to the success of this work going forward. The public must be confident that access will only be made available for legitimate research purposes to approved researchers and that the outputs from analysis will in no way compromise their privacy. The research publications should be publically available with any aggregate data and available for re-use under the Open Government Licence. The public must also be informed about the real benefits to citizens that will flow from the initiative in terms of better policy, better services and greater understanding of society and the economy.

As the structure to deliver the ADRCs is still under consideration, the plans for staffing and activities will become clearer as more specific proposals are developed. The role of public engagement and the debate around the use and reuse of linked de-identified administrative data will need to be co-ordinated across the network. Depending on the structure adopted for the ADRCs, this activity may best be provided through a central point or through the information gateway. The arrangements should be adequate for the purpose and should avoid duplication or lack of co-ordination but reflect and address the concerns and circumstances in the different parts of the UK.

R5 Sufficient funds should be put in place to support improved research access to and linkage between administrative data

R5.1 Funds should be secured to:

- establish the ADRCs, including technical and workforce requirements;
- support the data linkage activities of trusted third parties (organisations holding personal data which assist with secure data linkage procedures);
- support the development of the UK-wide researcher accreditation process and the provision of associated training courses;
- support the operational and strategic work of the Governing Board, including any legal work that is required, and the costs of the planned external review;
- support the independent auditing of the security procedures of the ADRCs;
- support for key activities (such as data retrieval, the creation of appropriate metadata, and the agreement on and implementation of agreed standards in data management) at the interface between the relevant government departments and the UK Administrative Data Research Network;
- assist Higher Education Institutions with the installation of secure rooms and the necessary equipment and staffing resource to provide virtual remote access to ADRCs.

R5.2 Funds to provide for the recommendations proposed in this report (comprising of the ADRCs, the governance structure, legal developments, accreditation and training, original research conducted within the ADRCs and the interface with government data holders) should be sought from an appropriate mix of interested parties including research councils, higher education funding councils, charities, and government.

Funding has been earmarked over the current spending review period for the investment to establish the necessary infrastructure to start developing ADRCs. The ESRC are currently developing a full business case which will go through the normal approval processes and government will be working with them to ensure that the points set out in this response are reflected adequately in implementation. Subject to approval of this business case, this funding will seed the infrastructure needed to develop capability and capacity to undertake analysis of linked data in a secure environment. We recognise that this activity will be incremental, particularly in the absence of enabling legislation.

The data owners remain responsible for the legal basis on which they allow access to their data so it is not anticipated that the ADRCs themselves would incur significant legal costs in relation to legalities of data release. Similarly, the basis on which partnerships, awards or investments are made with a variety of institutions should respect the need for a level playing field and avoid conferring competitive advantage. Where these involve Higher Education Institutions the investment should be consistent with the 2011 White Paper, *Higher Education: Students at the Heart of the System* ensuring no unfair advantage occurs.

The data of interest for this endeavour is held by government and part of the benefit of increased use of the data for research will accrue to government. The costs to government departments in supplying their data to the ADRCs may in some cases be substantial, particularly in the early development stages where new datasets are requested. However, there is already significant data made available by government departments through existing secure facilities such that there is likely to be data that can be made available to the research centres to start developing the necessary infrastructure and protocols at relatively low cost. The speed at which capacity and capability can be grown, along with the number, size and complexity of requests for data, will impact on the costs to data owners of releasing data.

It will be important as proposals are fine tuned and implemented that the implications for data owners are monitored. Other activities such as those supporting open data will also impact, in both an enabling and constraining way, on the capacity of government to respond to requests. There will be cases however where the data assets are stored in a way, or will require such work, that will make it prohibitively expensive to release at least in the short term. For some departments the benefits of being able to access the ADRCs will yield savings, for others the benefits may be less compelling and resources will need to be prioritised.

The establishment of ADRCs will mark a new departure for maximising the full potential of administrative data. The ambition is high and the challenges will be many but the expectation is that the benefits will be great. As the centres gain experience and expertise there will be ample scope for lesson learning in how data is collected, managed and used within Government which in turn will improve the statistical and analytical data holdings within Government and deliver better policies and better outcomes for citizens.

© Crown copyright 2013

You may re-use this information (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence. Visit www.nationalarchives.gov.uk/doc/open-government-licence, write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: psi@nationalarchives.gsi.gov.uk.

This publication available from www.gov.uk/bis

Any enquiries regarding this publication should be sent to:

Department for Business, Innovation and Skills
1 Victoria Street
London SW1H 0ET
Tel: 020 7215 5000

If you require this publication in an alternative format, email enquiries@bis.gsi.gov.uk, or call 020 7215 5000.

BIS/13/920