

Title: Introduction of a new power to allow public authorities to disclose de-identified data in controlled conditions for research in the public interest IA No: Lead department or agency: Cabinet Office Other departments or agencies: UKSA, HMRC, and DWP	Impact Assessment (IA)		
	Date: 01/02/2016		
	Stage: Development/Options		
	Source of intervention: Domestic		
	Type of measure: Primary legislation		
Contact for enquiries: Simon Meats (simon.meats@cabinetoffice.gov.uk)			

Summary: Intervention and Options	RPC Opinion: RPC Opinion Status
--	--

Cost of Preferred (or more likely) Option				
Total Net Present Value	Business Net Present Value	Net cost to business per year (EANCB on 2009 prices)	In scope of One-In, Two-Out?	Measure qualifies as
£m	£m	£m	Yes/No	In/Out/zero net cost

What is the problem under consideration? Why is government intervention necessary?

Data held by public authorities is of great potential value to researchers in government, academia, charities and industry. Access to more varied and better quality data would significantly improve the evidence base for research, including enabling more accurate analysis of economic and social issues to better inform policy design to improve public outcomes. But access has often been limited by a complex and uncertain legal landscape, resulting in research projects being delayed or abandoned. This proposal will provide a clear unambiguous power to allow public authorities to share data in safe settings for research in the public interest, as advocated by the research community for some time.

What are the policy objectives and the intended effects?

The policy objective is to facilitate a greater use of public data, encouraging a richer and more varied understanding of our economy and society. Specifically more data sharing for research purposes will assist in the delivery of a range of public benefits, from better policy design, to improved understanding of a economic and social issues and the better targeting of public services. We wish to introduce legislation to enable, across all public authorities (with exceptions) the sharing and linking of de-identified datasets for analysis by accredited researchers in a secure environment. The overall intention is to improve research outputs for public benefit whilst protecting privacy.

What policy options have been considered, including any alternatives to regulation? Please justify preferred option (further details in Evidence Base)

1) Option 1 – Do nothing: The status quo would be maintained, whereby specific statutory gateways are created when there is a need for them.
 2) Option 2 - Introduce general power for any and all public authorities to disclose identified data to each other for research purposes.
 3) Option 3 - Non-legislative work to change cultural boundaries: simpler guidance, brokerage of data-sharing agreements or other such provision.
 4) Option 4 (preferred option) - Introduce new legislation which enables the sharing of de-identified data by public authorities to accredited researchers using specified processes for research purposes in the public interest. This option reduces data sharing complexity and cost as it provides a clear purposive permissive power. It also balances this benefit with the protection afforded to the individual.

Will the policy be reviewed? It will not be reviewed. If applicable, set review date: Month/Year						
Does implementation go beyond minimum EU requirements?				Yes / No / N/A		
Are any of these organisations in scope? If Micros not exempted set out reason in Evidence Base.		Micro No	< 20 No	Small No	Medium No	Large No
What is the CO ₂ equivalent change in greenhouse gas emissions? (Million tonnes CO ₂ equivalent)				Traded: N/A		Non-traded: N/A

I have read the Impact Assessment and I am satisfied that, given the available evidence, it represents a reasonable view of the likely costs, benefits and impact of the leading options.

Signed by the responsible SELECT SIGNATORY: _____ Date: _____

Summary: Analysis & Evidence

Policy Option 1

Description: Do nothing

FULL ECONOMIC ASSESSMENT

Price Base Year	PV Base Year	Time Period Years	Net Benefit (Present Value (PV)) (£m)		
			Low: Optional	High: Optional	Best Estimate:

COSTS (£m)	Total Transition (Constant Price) Years	Average Annual (excl. Transition) (Constant Price)	Total Cost (Present Value)
Low	Optional	Optional	Optional
High	Optional	Optional	Optional
Best Estimate			

Description and scale of key monetised costs by 'main affected groups'

In line with impact assessment guidance the do nothing option has zero costs or benefits as impacts are assessed as marginal changes against the do nothing baseline.

Other key non-monetised costs by 'main affected groups'

In line with impact assessment guidance the do nothing option has zero costs or benefits as impacts are assessed as marginal changes against the do nothing baseline.

BENEFITS (£m)	Total Transition (Constant Price) Years	Average Annual (excl. Transition) (Constant Price)	Total Benefit (Present Value)
Low	Optional	Optional	Optional
High	Optional	Optional	Optional
Best Estimate			

Description and scale of key monetised benefits by 'main affected groups'

In line with impact assessment guidance the do nothing option has zero costs or benefits as impacts are assessed as marginal changes against the do nothing baseline.

Other key non-monetised benefits by 'main affected groups'

In line with impact assessment guidance the do nothing option has zero costs or benefits as impacts are assessed as marginal changes against the do nothing baseline.

Key assumptions/sensitivities/risks	Discount rate	
-------------------------------------	---------------	--

BUSINESS ASSESSMENT (Option 1)

Direct impact on business (Equivalent Annual) £m:	In scope of OITO?	Measure qualifies as
Costs:	Yes/No	IN/OUT/Zero net cost
Benefits:		
Net:		

Summary: Analysis & Evidence

Policy Option 2

Description: Introduce a general power for all public authorities to disclose identified data to each other for research purposes

FULL ECONOMIC ASSESSMENT

Price Base Year 2015	PV Base Year 2015	Time Period Years 10	Net Benefit (Present Value (PV)) (£m)		
			Low: Optional	High: Optional	Best Estimate: 0

COSTS (£m)	Total Transition (Constant Price) Years	Average Annual (excl. Transition) (Constant Price)	Total Cost (Present Value)
Low	Optional	Optional	Optional
High	Optional	Optional	Optional
Best Estimate	0	0	0

Description and scale of key monetised costs by 'main affected groups'

N/A

Other key non-monetised costs by 'main affected groups'

It is expected that public sector bodies affected by the legislative change will face one-off familiarisation and training costs associated with the change in legislation. Public sector bodies will also incur administrative costs associated with an increased volume of data sharing requests. However, it is expected that any data sharing burden would be at least partially offset by benefits associated with increasing the quality and quantity of policy relevant research (see below).

BENEFITS (£m)	Total Transition (Constant Price) Years	Average Annual (excl. Transition) (Constant Price)	Total Benefit (Present Value)
Low	Optional	Optional	Optional
High	Optional	Optional	Optional
Best Estimate	0	0	0

Description and scale of key monetised benefits by 'main affected groups'

N/A

Other key non-monetised benefits by 'main affected groups'

Public sector bodies will benefit from a decrease in the administrative costs of sharing data (i.e. staff time in researching or establishing legal data sharing gateways). The public sector will be better able to create government services that draw upon the more sophisticated understanding of patterns in users and behaviours to tailor services more closely to the individual, improving quality and reducing costs - enabled through public sector bodies' ability to share data more quickly.

Key assumptions/sensitivities/risks

Discount rate

The key risk of a legislative change lies in the possibility of future legal challenge with respect to the Data Protection Act or the Human Rights Act. This is most significant under this option as it enables a wide range of data sharing with identified data. Related to this, there is a risk in data loss and associated personal costs to citizens as well as reduced trust in government. A further risk is in incorrect use of data.

BUSINESS ASSESSMENT (Option 2)

Direct impact on business (Equivalent Annual) £m:			In scope of OITO?	Measure qualifies as
Costs:	Benefits:	Net:		

Summary: Analysis & Evidence

Policy Option 3

Description: Non-legislative work to change the culture around data sharing in the public sector

FULL ECONOMIC ASSESSMENT

Price Base Year 2015	PV Base Year 2015	Time Period Years 10	Net Benefit (Present Value (PV)) (£m)		
			Low: Optional	High: Optional	Best Estimate: 0

COSTS (£m)	Total Transition (Constant Price) Years	Average Annual (excl. Transition) (Constant Price)	Total Cost (Present Value)
Low	Optional	Optional	Optional
High	Optional	Optional	Optional
Best Estimate	0	0	0

Description and scale of key monetised costs by 'main affected groups'

N/A

Other key non-monetised costs by 'main affected groups'

Public sector bodies will face one-off costs related to the development and implementation of a culture change/training programme. This would involve training delivered to staff who are or could be involved in data sharing. Assuming a successful shift to increasing the level data sharing within the existing legal framework, public sector bodies will also face ongoing administrative costs associated with sharing data, but it is expected that these would be at least partially offset by benefits associated with increasing the quality and quantity of policy relevant research.

BENEFITS (£m)	Total Transition (Constant Price) Years	Average Annual (excl. Transition) (Constant Price)	Total Benefit (Present Value)
Low	Optional	Optional	Optional
High	Optional	Optional	Optional
Best Estimate	0	0	0

Description and scale of key monetised benefits by 'main affected groups'

Other key non-monetised benefits by 'main affected groups'

Assuming a successful shift to increase data sharing within the existing legal framework, public sector bodies will benefit from a decrease in the administrative costs of sharing data (i.e. staff time in researching or establishing legal data sharing gateways where gateways already exist). The public sector will be better able to create government services that draw upon the more sophisticated understanding of patterns in users and behaviours to tailor services more closely to the individual - enabled through public sector bodies' ability to share data more quickly. The magnitude of these benefits is expected to be lower than under options 2 and 4.

Key assumptions/sensitivities/risks

Discount rate

BUSINESS ASSESSMENT (Option 3)

Direct impact on business (Equivalent Annual) £m:			In scope of OITO?	Measure qualifies as
Costs:	Benefits:	Net:	No	NA

Summary: Analysis & Evidence

Policy Option 4

Description: Introduce legislation which enables the sharing of de-identified data by public authorities to accredited researchers for research purposed in the public interest

FULL ECONOMIC ASSESSMENT

Price Base Year 2015	PV Base Year 2015	Time Period Years 10	Net Benefit (Present Value (PV)) (£m)		
			Low: Optional	High: Optional	Best Estimate: 0

COSTS (£m)	Total Transition (Constant Price) Years	Average Annual (excl. Transition) (Constant Price)	Total Cost (Present Value)
Low	Optional	Optional	Optional
High	Optional	Optional	Optional
Best Estimate	0	0	0

Description and scale of key monetised costs by 'main affected groups'

Other key non-monetised costs by 'main affected groups'

It is expected that public sector bodies affected by the legislative change will face one-off familiarisation and training costs associated with the change in legislation. Public sector bodies will also incur administrative costs associated with an increased volume of data sharing requests. There will be further administrative costs to the UKSA who will act as an accrediting body to those who want to access data, and to the Trusted Third party indexer, who will facilitate the matching of data. Again, it is expected that any data sharing burden would be at least partially offset by benefits associated with increased research in the public interest.

BENEFITS (£m)	Total Transition (Constant Price) Years	Average Annual (excl. Transition) (Constant Price)	Total Benefit (Present Value)
Low	Optional	Optional	Optional
High	Optional	Optional	Optional
Best Estimate	0	0	0

Description and scale of key monetised benefits by 'main affected groups'

Other key non-monetised benefits by 'main affected groups'

Public sector bodies will benefit from a decrease in the administrative costs of sharing data (i.e. staff time in researching or establishing legal data sharing gateways). The public sector will be better able to create government services that draw upon the more sophisticated understanding of patterns in users and behaviours to tailor services more closely to the individual, improving quality and reducing costs - enabled through public sector bodies' ability to share data more quickly.

Key assumptions/sensitivities/risks

Discount rate

The key risk of a legislative change lies in the possibility of future legal challenge with respect to the Data Protection Act or the Human Rights Act. Related to this, there is a risk in data loss and associated personal costs to citizens as well as reduced trust in government. A further risk is in incorrect use of data in policy-making.

BUSINESS ASSESSMENT (Option 4)

Direct impact on business (Equivalent Annual) £m:			In scope of OITO?	Measure qualifies as
Costs:	Benefits:	Net:		

Evidence Base (for summary sheets)

Problem under consideration

Data held by public authorities is of great potential value to researchers in government, academia, charities and industry. Access to more varied and better quality data helps to create an improved evidence base for researchers both within and outside government, enabling better informed analysis and research, a deeper, more granular understanding of what underpins key policy challenges, in order to improve the evidence base for policy development to improve a range of public benefits. Linking datasets held by two or more public authorities in a controlled environment offers the opportunity to gain new insights into the social and economic challenges that citizens face. It would provide better understanding of how people live their lives, the patterns of need and use of different services, and the resultant outcomes. It would allow the delivery of better, more efficient and more effective services and processes, leading to improved understanding and responses to challenges relating to the health and wellbeing of citizens. Currently in the majority of cases, research using government-held data is limited to the analysis of single data sets. Consequently the possibility to undertake deeper research using cross-linked but separate datasets is both rare and difficult.

At present researchers are often frustrated in their efforts to access public sector-held data for research projects that have a potential public benefit. The current legislation causes public authorities to be uncertain as to what information can be disclosed. The issue of whether disclosing a particular dataset is lawful can lead to lengthy delays and inconsistent decisions around access. Meanwhile, the possible economic or social benefit from a proposed research project could be lost due to delays in reaching a decision, with the risk of some potentially valuable projects being abandoned. It has been the view of representatives of the research and statistics community for some time that this situation is overdue for reform.

Given the sensitivities around data sharing issues, it is essential that any regime for linking publicly-held datasets for interrogation by researchers must be:

- secure in its operations,
- transparent in its governance, and
- in the public interest

Examples of frustrated or seriously delayed research projects include:

- To consider UK productivity and growth and understand the productivity gap, linking data from a number departments including their earning and social welfare status g DfE, HMRC, and DWP to explore relationships between individuals' employment, qualifications and the profitability of business
- Linking data relating to higher education, employment/benefits and earnings to explore the relationship between individuals moving between secondary or higher education and the labour market, including variables such as receiving free school meals or working age benefits
- The Chief Medical Officer for Wales' to access housing data from the Valuation Office Agency to explore the relationship between winter mortality and housing conditions
- Research to link earnings and employment data to explore the relationship between social mobility and age cohorts

- A detailed analysis of how exposure to international trade and the export activities of UK firms impacts upon firm-level indicators such as productivity or innovation

Further detail on the legal landscape

In researching this problem, Cabinet Office surveyed a variety of public bodies. The results of this were analysed by CO policy lawyers. This revealed that many could use the models of data sharing proposed, but they could not all do so in all circumstances. Statutory Bodies for example, and for this policy most importantly HMRC, are severely curtailed as regards what they can do as their vires is strictly limited to what is specifically set out in statute. Some bodies may have either explicit data sharing powers or alternatively may be interpreted as being able to conduct data sharing through the use of some general ancillary power, such as Local Authorities, who have relied on the power to promote or improve social, economic or environmental well-being in Section 2 of the Local Government Act 2000, and since February 2012, on the general power of competence in s1 the Localism Act 2011. However where such powers are lacking or insufficient, additional legislative provision will need to be provided by our legislation to enable them to use the models of data-sharing as set out in this document for the purposes of accredited research by accredited researchers as set out in this document.

In contrast Ministers and Government Departments for example are able to rely on common law powers under the principles set out in the Ram doctrine. Therefore all other things being equal, such persons may in theory be able to rely on common law powers as a legal person to engage in data sharing. However for the purpose of this policy most crucially DWP have told us that its ability of to rely on these common law powers in respect of DWP data has been eroded by the passing of legislation that governs the same area. The number of gateways for DWP data is claimed by DWP as the reason its powers under Ram doctrine have been eroded.

For these reasons, some Departments (and for the purpose of this policy in particular HMRC and DWP) are often not willing to engage in data sharing for research, and it is data from HMRC and DWP that is most demanded by researchers that we have consulted. In addition other bodies lack clarity as to whether such a power exists. University and third sector researchers have confirmed to us that this lack of clarity is often given to them by departments as a reason as to why they are unable to be provided with linked de-identified administrative data.

Rationale for intervention

Wider use of data sharing would facilitate an improvement in understanding into our economy and society, and provide a more informative platform for policy development. More informed evidence based policy will encourage the creation of government services that draw upon the more sophisticated understanding of patterns in users and behaviours to tailor services more closely to the individual to improve quality and reduce costs; it will enable more efficient timing of interventions and make policy evaluations easier.

By restricting access to data not by intent but by the arbitrary and complex pattern of primary and secondary legislation that has often followed the creation of Government departments and agencies, we risk stemming the potential benefits of that data to research and by extension to our economy and society. At present, research projects are subject to delays of up to two years whilst a bespoke legal gateway is created.

Introducing the proposed new legislation will improve conditions for research in two major respects:

- Public authorities will have much greater clarity about what data is permitted to be shared.

- The proposed power places conditions on the secure disclosure of data to provide additional assurance to public authorities, researchers and the public that their data is being used correctly. It requires the use of specific safeguards to ensure that any information that could be used to identify, or help to identify, an individual (e.g. names, date of birth and postcode) is de-identified through privacy-enhancing processes.

The need for intervention was highlighted by the December 2012 report from the Administrative Data Taskforce (ADT)¹. This was formed in December 2011 with the aim of improving access to and linkage between government administrative data for research purposes for researchers both inside and outside government. A key recommendation of this report was the creation of a new generic gateway to allow access to publicly held data by the research community. The ADT report also recommended some models of privacy enhancing data sharing, to allow data to be linked in a secure way. The legislation attempts to provide powers for these models to be used, where such powers do not already exist.

The case for new legislation, including specifying the use of established process models to protect personal privacy, was explored fully during the open policy-making discussions in 2014-15, and it was agreed that access to data that has been de-identified for research purposes under this power would be conditional on the use of such methods. The open policy-making group, which included representatives from organisation involved in or connected with the ADT report, reached consensus on this and other proposals for linking and de-identifying data for research purposes.

Policy objective

To enable conditions that facilitate more data-sharing for research that will assist in the delivery of a range of public benefits, from better policy design, to improved understanding of a variety of economic and social issues and the better targeting of public services.

The policy is not to make access mandatory, but just to make it possible, if the data owner wishes it. Whilst there is no guarantee that outside researchers will obtain greater access to data as a consequence of these provisions as access will still be dependent on the permission of the relevant data controller. However it is hoped by making the legislative and administrative provisions to enable these models of data sharing, that overall access will increase, leading to more sharing taking place due to public authorities and research organisations being 'comfortable' to share data where there is an established model to follow. This is important for the following reasons:

- i) This will allow for a richer analysis to be produced from a wider range of variables leading to greater flexibility for analytical/research programmes;
- ii) It would provide better understanding of how people live their lives, the patterns of need and use of different services, and the resultant outcomes. This will allow for the delivery of better, more efficient and more effective services and processes leading to better outcomes for the UK Government and the public.

A supporting objective is to provide a secure framework for public authorities to allow the research community access to data. This framework requires the use of established process models to protect personal data including:

¹ <http://www.esrc.ac.uk/collaboration/collaborative-research/adt/>

- The removal of any information that might be used to identify an individual or organisation contained in a dataset and its replacement by reference numbers by the data holder (thus “de-identifying” the data before it is sent for processing)
- The use of a “Trusted Third Party (TTP)” indexer to allow the matching of two or more datasets without disclosing identifying information to researchers
- The use of an accredited access facility to link two or more de-identified datasets for analysis by researchers in controlled conditions

In addition, entry to the regime by researchers and those providing indexing or secure access facilities will be subjected to accreditation by a designated body, which has been specified as the UK Statistics Authority. The UKSA will be responsible for setting guidance, including parameters as to what may be considered to be research in the public interest, and will have the power to set the condition for accreditation, and for withdrawing accreditation. It will also have certain responsibilities regarding transparency, including publishing a register of accredited researchers and the type of research being undertaken. This will support the need to provide public assurance.

Who the policy is meant to apply to

- All public authorities (with the exception of public bodies providing health and social care). We are defining a public authority as a person with functions of a public nature (consistent with the definition in the Human Rights Act)
- The UK Statistics Authority, as the designated accreditation body
- Those research organisations and individual researchers seeking accreditation under the new power, who wish to participate
- Providers of secure access and TTP facilities, including government departments with in-house facilities such as HMRC’s datalab, who wish to participate

Exclusions – health bodies

Certain organisations and persons should not be able to benefit from the extension in the vires. These are as follows:

- a. specifically defined bodies delivering health services² and adult social care³ should be excluded;
- b. any person who provides health services, or adult social care, pursuant to arrangements made with a public body exercising functions in connection with the provision of such services or care should be excluded.

² “Health services” means services which must or may be provided as part of the health service as defined by s275(1) of the NHS Act 2006

³ **Adult social care** includes all forms of personal care and other practical assistance provided for individuals who, by reason of age, illness, disability, pregnancy, childbirth, dependence on alcohol or drugs, or any other similar circumstances, are in need of such care or other assistance

Private health and care service providers are not to be specifically excluded as they are not public bodies, and therefore would not be affected by any extension in the vires.

The carve-out for bodies delivering health services is not a prohibition. It is merely an carve-out from any enhancement in their vires, and therefore any public body delivering health services that has sufficient vires already to use the models will be unaffected by the exclusion.

It is not health data per se that is excluded, since DWP holds health data that they may have received from the health bodies, such as the disability status of an individual. Such health data in the hands of DWP should be able to be disclosed by DWP to accredited researchers under the model.

The justification for the health and social care bodies data exclusion (which has been provided by Department of Health officials after consultation with their lawyers) is that clinicians, patients and members of the public in England have all expressed serious misgivings about sharing health information for secondary purposes – even where these purposes are healthcare related. The recent controversy over “care.data” – a pseudonymised dataset containing person-level primary care data – has highlighted these concerns and the strength of feeling. Implementation of care.data is currently on hold while the Department of Health consults with stakeholders over the additional safeguards required to reassure patients and clinicians that confidential data will not be misused.

Similar exclusions to those described above will have to be constructed for health bodies in the devolved administrations to ensure consistency of approach across the UK.

Other exclusions

- Processing of personal data under the power must be for research purposes. Though the outputs of a research project under the power can be used to inform strategic decisions on policy and operations, the power cannot be used for specific decisions against data subjects (i.e. such as the loss of a benefit). The policy developed in partnership with civil society groups, is intended to ensure that data shared for research purposes is not directly used to the detriment of a data subject(s).

Options considered

The following options were identified and appraised:

- 1) Option 1 – Do nothing: The status quo would be maintained, whereby specific statutory gateways are created when there is a need for them.
- 2) Option 2 - Introduce general power for any and all public authorities to disclose identified data to each other for research purposes.
- 3) Option 3 - Non-legislative work to change cultural boundaries: simpler guidance, brokerage of data-sharing agreements or other such provision.
- 4) Option 4 (preferred option) - Introduce new legislation which enables the sharing of de-identified data by public authorities to accredited researchers using specified processes for research purposes in the public interest.
 - a. a broad generic power: a general power for any and all public authorities to disclose identified data to each other for research and statistics; at the Open Policy Making Plenary meeting on 22 October 2014, members of Civil Society, present in person and by proxy, made it clear that they felt that this option would be a step too far.
 - b. a power for a single data controller to use a safe haven to process or disclose their own data for research or statistical analysis;
 - c. do nothing; and

- d. a power for public authorities (except health services bodies and adult social care bodies that are specifically defined) to link de-identified data for research using accredited bodies and TTP sharing (recommended approach - see below).

Description of Options considered

Option 1: Do Nothing: Allowing the creation of a number of specific statutory gateways, where there is a need for them.

This option is essentially about maintaining the current status quo and making provision for each data-share required for the purposes of research, through existing individual statutory gateways or by creating new statutory gateways.

Costs: The cost associated with creating specific legal gateways is significant, in terms of official, lawyer and Parliamentary time spent firstly in understanding the complex legal landscape, creating an appropriate gateway suitable for the specific needs and then passing the gateway through Parliament. Some legislative barriers will require new primary legislation to overcome them, which is subject to even greater competition for Ministerial or Parliamentary time.

The costs to the research community and the wider public interest will be that the problems outlined earlier in this document will remain. This would not provide the research and policy benefits that will derive from making available a more transparent, more consistent method of linking de-identified datasets for any research purpose.

Benefits: Specific gateways (created through secondary legislation) are least burdensome when it comes to producing evidence that such a gateway is required; in order to be added to statute it will be subject to Parliamentary scrutiny and therefore deemed necessary. To do so, it is expected that the case for the gateway will be clearly set out, providing assurance that such a measure was proportionate and necessary before it was created.

Option 2 - Introduce general power for any and all public authorities to disclose identified data to each other for research purposes.

As a permissive power, this option would not place any statutory obligations on public authorities to share data, but would allow them to share data for research purposes if two or more agreed to do so in any given case. It would not require the process models and accreditation process referred to earlier, as this would involve public authorities sharing between themselves (and presumably producing de-identified outputs if they wished). It would also require a robust set of safeguards to protect personal data. Further, it does not allow access to publicly-held data by research groups; this was rejected by civil society organisations during the open policy-making process, who felt that it went too far in terms of risks to personal privacy.

Costs: Being a permissive power, it is hard to estimate costs for this option as they are largely dependent on uptake by public authorities. For those authorities who wished to share data, it is expected that one-off costs would apply in respect of the need to familiarise officials with the new legislation. These may include time resource for policy and legal officials in researching/disseminating information, and training.

Given that this power could potentially involve the sharing of identified data across the public sector, it is expected that certain standards for the protection of personal data would be part of the conditions specified (either in primary or secondary legislation, or through a code of practice or similar). These would include certain technical standards or practices to encrypt data that would be processed under this power, together with the compiling and publication of privacy impact assessments, and compliance with

existing legislation including the Data Protection Act (DPA). Whilst compliance with these standards would be less likely to create new burdens to those authorities who are familiar with data sharing arrangements, some one-off costs may apply to those organisations (particularly local authorities), who have less developed approaches to information governance in complying with new guidance. Therefore there would be both one-off costs by authorities in familiarising/levelling up of standards, and on-going costs applying to all authorities who agreed to share data under this power in respect of complying with the new standards. Continuing costs would also apply in respect of staff time resource in compiling and processing data that is requested.

There would also be one-off costs applicable to whichever authority was responsible with creating guidance/a code of practice (policy and legal officials, consultation with Information Commissioner etc), together with the creation of some form of enforcement mechanism. On-going costs would then apply in terms of maintaining and enforcing guidance/codes of practice.

Benefits: We would expect an administrative saving from the reduction in the number of future data sharing agreements between Departments that would need to be set up individually compared with the hypothetical 'do nothing' scenario.

Overall, we might expect the net impact on public authorities to be a net saving as increases in the volume of data sharing as a result of the legislation is expected by public authorities to be outweighed by a reduction in the cost of sharing each element of data, and in setting up individual agreements. Non-monetised benefits both to the public sector and to the economy and society would be gained from maximising the opportunities for sharing data between public authorities and increase the amount of policy-relevant research that resulted. This would maximise the data available to assist in policy design and evaluation, and in turn create the opportunities for more innovative, efficient and better targeted public policy development.

Option 3 - Non-legislative work to change cultural boundaries: simpler guidance, brokerage of data-sharing agreements or other such provision.

This option recognises the issue that a number of the barriers to effective data-sharing are cultural: overly-cautious interpretation of statute, trust in how other organisations will use data, incentives to withhold the supply of data, a lack of confidence in the integrity of the data being shared and a lack of consistent standards in definitions, formats and collection methodology. All these add to the issue of public authorities being reluctant to share with each other and more broadly.

Some work has been undertaken, for example through the establishment of the Administrative Data Research Network (ADRN), the result of another recommendation from the ADT report, towards assisting the scientific community in facilitating a secure environment for data access in safe settings. The Government also provided funding to support the ADRN's establishment, although its representatives claim that further funding is needed. However, cultural change and funding alone will not solve the problem of statutory barriers or uncertainty that has prevented public authorities from sharing data.

Costs: There will be a cost in terms of creating, disseminating and training staff with the new guidance. However, this does not reduce *actual* legislative complexity, or the time taken, or the flexibility of any legislation of itself. As legislation would still be required, there would still be Parliamentary time and Administrative costs of setting up agreement when needed, as in the "do nothing" case.

Benefits: This option would reduce *perceived* complexity of the legislation as clearer guidance is developed and best practice shared which will reduce *cultural* barriers to data sharing. As stated above, they will not reduce actual complexity or remove barriers.

Option 4 (preferred option) - Introduce new legislation which enables the sharing of de-identified data by public authorities to accredited researchers using specified processes for research purposes in the public interest.

This will be permissive power to ensure that public bodies (except health services bodies and data held by bodies in respect of their role in adult social care) are able, if they so wish, to engage, for the purposes of research or statistical analysis, in the process of linking two or more datasets from two or more data controllers using a secure method of sharing (referred to earlier as the Trusted Third Party model).

This option reduces data sharing complexity and cost both to public authorities and researchers who use this power as it provides a clear purposive permissive gateway. It also balances this benefit with the protection afforded to the individual.

Costs: This is a permissive power, therefore determining costs would largely be dependent on the level of uptake by public authorities, as for Option 2.

In terms of costs to public authorities, there would be both one-off costs by authorities in familiarising/levelling up of standards, and on-going costs applying to all authorities who agreed to share data under this power in respect of complying with the new standards. Continuing costs would also apply in respect of staff time resource in compiling and processing data that is requested.

In terms of cost burdens on the UKSA, as the designated accrediting body UKSA will have an administrative cost burden. We have outlined the main responsibilities here and will develop a more detailed understanding of the cost during the consultation period. The main activities required to be undertaken will include:

- Establishing a new system for accreditation and developing a Code of Practice through consultation with the Minister for the Cabinet Office, the Information Commissioner and any other persons as deemed appropriate. This Code would need to be supported by a set of technical standards, which would in turn need to be kept up-to-date
- Publication of new system and invitations to research and statistics community; for example through a website, which would need managing and updating
- Processing applications to participate under this system, ensuring smooth passage of the application and communicating decisions to the applicant. An appeals process for rejected applications might also be necessary. Final decision/oversight of process would be made through a panel/board which considers recommendations from officials
- Maintain and publish:
 - A list of accredited bodies who facilitate research, as Trusted Third Parties, secure access facilities etc
 - A list of organisations, research projects and individual level researchers who are permitted to undertake researcher under this system.
 - Information regarding those applications that have been rejected (format to be agreed with UKSA).
- Enforcing Code and sanctions for breaches, the accrediting body will need to establish a mechanism to enforce sanctions against breaches of the code.
- Continuous improvement and review of standards

This proposal does not provide for any charging mechanism for a researcher seeking accreditation, and we understand that there is no current intention for secure access facilities to charge for accreditation or to charge for access to data.

There are additional costs of using a Trusted Third Party Indexer. The Trusted Third Party Indexer is body or person accredited under the system set up in the proposed legislation, which would receive only the "identity dataset" extracted from the data sets of two or more data owning authorities that will be matched, and who is then tasked with matching up records in each set so as to be able to produce a linkage document or a linking reference, the "index." This index is then passed to the accredited access facility, allowing it, for each individual data subject, to match up the de-identified data from one dataset with the specific corresponding de-identified data from every other dataset involved.

A TTP could consist of a single analyst with the necessary indexing capability or a data laboratory; however it would need to conform to standards set by the UKSA as the accrediting body. As TTP Indexing is already current practice (in for example the Administrative Data Research Network), we

would not expect the proposed legislation to call for new technology to be developed. However, best practices and standards would need to be adhered to.

Those organisations who provide TTP facilities (including government departments) may, depending on the uptake of the proposal (it being a permissive power), experience increasing demands on their services, possibly requiring more data analysts or facilities to be employed/deployed.

Benefits: The monetised benefits of the proposal are not possible to quantify, as it is not possible to predict the increase in volume of data shares arising from the use of this power. However, key non-monetised benefits arising from the proposal would include the following:

It would remove a layer of legal barriers and uncertainties which have caused considerable delays and have in some cases prevented worthwhile research projects from going ahead, as referenced earlier. This would enable savings to public authorities in terms of time resource for policy and legal officials in determining whether a particular proposal was lawful.

It would enable research leading to better-informed policy in areas such as:

- Address social mobility, by linking data on employment, training, education, unemployment, incomes and benefits;
- Enable research of causal pathways over the life course – linking data on education, employment, incomes and wealth;
- Inform policies designed to tackle poverty – linking data on housing conditions, incomes, and benefits
- Construct indicators of parental employment, social background, and childcare.
- Link data on offending behaviour, incomes and benefits.

It is expected that better informed policy will encourage the creation of government services that draw upon the more sophisticated understanding of patterns in users and behaviours to tailor services more closely to the individual to improve quality and reduce costs; it will enable more efficient timing of interventions and make policy evaluations easier.

Risk and Assumptions

The proposed changes are intended to improve public sector bodies' ability to share de-identified data with researchers inside and outside of Government to improve policy design and public outcomes. The risks that these changes will bring about are common to any data sharing process, namely:

- a) Loss of data;
- b) Incorrect use of data – with biased or incorrect conclusions being drawn and policy ineffectively designed as a result;
- c) Challenge from individuals whose data has been shared.

The use of data sharing has increased substantially in recent years and it is encouraged within Government to make better use of existing information. This has meant a better understanding of the risks associated with it. As a result, a number of measures have been developed to mitigate these risks. These mitigation measures are either required by law or considered as good practice and include among others:

- Organisations sharing data have the appropriate organisational measures in place as established by the Data Protection Act. It is good practice to:
 - design and organise security to fit the type of personal data disclosed or received and the harm that may result from a security breach
 - be clear about which staff members in the organisations involved in the sharing are responsible for ensuring information security

- have an appropriate monitoring and auditing procedure in place
- be ready to respond to any failure to adhere to a data sharing agreement swiftly and effectively
- Organisations sharing data have the appropriate technical measures in place as established by the Data Protection Act. It is good practice to:
 - make sure that the format of the data you share is compatible with the systems used by both organisations
 - check that the information that is shared is accurate before sharing it
 - establish ways for making sure inaccurate data is corrected by all the organisations holding it
 - agree common retention periods and deletion arrangements for the shared data
 - train staff so that they know who has the authority to share personal data, and in what circumstances this can take place.
- The various organisations involved in data sharing will each have their own responsibilities and liabilities in respect of the data they disclose or have received. It is therefore good practice:
 - for a senior, experienced person in each of the organisations involved in the sharing to take on overall responsibility for information governance, ensuring compliance with the law, and providing advice to staff faced with making decisions about data sharing
 - to have a data sharing agreement in place that includes:
 - The purpose of the sharing
 - The potential recipients or types of recipient and the circumstances in which they will have access
 - The data to be shared
 - Data quality – accuracy, relevance, usability, etc
 - Data security
 - Retention of shared data
 - Individual's rights – procedures for dealing with access requests, queries and complaints
 - Review of effectiveness/termination of the sharing agreement, and
 - Sanctions for failure to comply with the agreement or breaches by individual staff.

Overall, the appropriate mitigating measures depend on the type of information that is shared and the organisations that are sharing them. Therefore, any future policy that requires the use of data sharing should specify what mitigating measures are more appropriate to reduce risks.

Annex A

Detailed description of the process

Overview

Data legislative measures will provide all UK public bodies who 'would not otherwise have power to make the disclosure', with a power, where the purpose of the share or disclosure is to enable processing for research purposes, and the data cannot be further shared other than with the consent of the sources and can only be disclosed to accredited persons to:

- a) Disclose both personal data and non-personal data to an accredited external data processor for the purpose of indexing against another dataset. In a similar process model, the Indexer also needs to ability to notify where there have been matches to the data sources. This means that the Indexer can generate the reference number for matching identifiers, and send those back to data-holding authorities. These authorities then attach the reference numbers to the payload data and send those to the secure accredited facility. In this model there is no communication between the Indexer and the Secure Access Facility.
- b) Share de-identified personal and normal data as well as indexes with secure access facility.
- c) Share de-identified personal and normal data with accredited researchers and analysts (as external data controllers, who are either public or private persons) or alternatively disclose de-identified personal and normal data to researchers and analysts (where they act as a data processor, who are either public or private persons).

Our proposals are intended to enable these variant models of data sharing not in the sense that the Bill will implicitly or explicitly repeal all contrary law and guarantee that such data shares can take place. Instead, the proposed legislation should enable these variant data share models in the sense that it will remove the initial constraints that prohibit public bodies from even getting to the stage of considering if a data share would comply the DPA and HRA. The reason for this is that these constraints are the barrier to use of the models that has stopped HMRC and DWP from using them.

Other conditions set out in legislation

The essential elements of the process being legislated for are that:

- a) there will be a process that the data owners use to match the identifiers using an accredited indexer before the data owners make de-identified data available in an secure access facility to an accredited researcher for accredited research;
- b) the data sources are empowered to pass the identity dataset to an accredited indexer for the purposes of (i) matching and (ii) creating an index;
- c) the indexer cannot be one of the data sources; this is privacy enhancing;
- d) the indexer is empowered to pass the index to the secure accredited environment;
- e) the indexer is empowered to provide the identifiers of those individuals who have been matched back to the data sources together with any study-identifiers;
- f) the indexer is empowered to provide the identifiers of those individuals who have not been matched back to the data sources;
- g) the data sources are empowered to provide the payload datasets to the secure access facility for the purpose of linking them and making them available to accredited researchers for accredited research.
- h) the secure access facility cannot be one of the data sources – this is necessary because otherwise, where DWP and MOJ were wanting to link the payload data they held, as an secure access facility for that data share, DWP would have access to all the payload data of MOJ and DWP;
- i) the public authorities who use the vires power, the secure access facility, and the Indexer may charge on a cost recovery basis for their services. If the Indexer is not a public authority it should be able to charge a commercial rate for its services.

The above process will be set out in the proposed legislation allowing data sharing only under these conditions.

Safeguards and accreditation

Given that the data sharing power above is wide in both the scope of bodies it would apply to, and the scope of material covered, it is appropriate that the legislation would also need to include specific safeguard provisions. Therefore, the vires provision is made subject to a condition that it may only be used when all the bodies and individuals involved in a data share (other than the data sources) are accredited bodies. The legislation therefore also specifies that the UK Statistics Authority will be the accreditation body on the basis that exercises functions throughout the whole of the UK, and it has the necessary expertise in research and statistical analysis.

The accreditation body would accredit the indexers, the researchers, the secure access facilities and the research itself. Minimum accreditation requirements are to be set out in the legislation as an important safeguard to reassure the privacy lobby and to mitigate the risk of re-identification. The accreditation body would themselves develop and publish additional detailed standards and requirements to attain and maintain this accreditation.

The data that is processed and shared using the proposed power is separated into payload data and identifiers. Therefore it is important that there are safeguards against the risk of re-identification by the researchers and secure access facilities. In order to provide the safeguards, the accreditation body needs to be provided by the primary legislation with a power to accredit under primary legislation. They would be expected to accredit the following:

- Secure access facilities
- Indexers
- Researchers
- Research

The legislation will require the UKSA to keep a central register of those who are accredited, indicating in each case the specific accreditation(s) that they hold. They must publish this register (preferably online) and keep it current.

Where the accrediting body becomes aware that an accredited person is for some reason no longer considered to be a fit and proper person in the light of published criteria, the accrediting body must be able to remove the accreditation for the purposes of future data shares. The appeals process for removal of accreditation is the same as that for refusal of an initial grant of accreditation and is described later in this document.

The minimum standard for those who are to be accredited to meet is set out in the primary legislation in order to provide privacy assurance. The minimum standard is "Fit and proper". The accrediting body must from time to time publish criteria by reference to which it will determine whether to grant accreditation; it must consult on these criteria before publication.

Primary legislation will specify that secure access facilities:

- Must be fit and proper;
- Must set conditions to access the service of the secure access facility;
- The accrediting body must from time to time publish criteria by reference to which it will determine whether to grant accreditation; and
- Their identities to be published by the accrediting body.

In practice those bodies likely to seek and be granted accreditation as secure access facilities are individual Universities, Research Institutes and Statistical Institutes who will receive the de-identified data from the data sources. The ESRC (Economic and Social Research Council, who sponsor the work of the ADRN) has told us that the current practice is that the Universities who hold the administrative data in their safe havens are acting as Data Processors. These Universities process the data on behalf

of the data controller (the source department) but do not “determine the purposes for which and the manner in which any personal data are, or are to be, processed.”

Primary legislation will specify that an Accredited Indexer:

- must be fit and proper; the accrediting body must from time to time publish criteria by reference to which it will determine whether to grant accreditation;
- the identity of those persons/bodies that have Accredited Indexer status must be published on the website of the accrediting body; and
- Accredited Indexers do not have to be public authorities.

Primary legislation will specify that an Accredited TTP Researcher:

- must be a fit and proper person; the accrediting body must from time to time publish criteria by reference to which it will determine whether to grant accreditation; as a matter of precedent the SRSA contains provisions for “approved researcher status” in s39(5-6). There should be a provision to oblige the accrediting body to consult on these criteria.
- conducting accredited research; as a matter of policy we do not intend to exclude the possibility of private bodies or persons becoming accredited researchers.

The researchers will, without specific provision in the legislation, be subject to all the existing law that applies to the data that they are accessing e.g. statutory bars on disclosure contained in Statistics and Registration Services Act, Commissioners for Revenue and Customs Act 2005, the requirements of Data Protection Act, the law of confidence, the contractual conditions of access set by the data source controllers, which the researchers will sign.

Primary legislation will specify that Accredited TTP research:

- research that is, in the opinion of the accrediting body or any body to which it validly sub-delegates its power, in the public interest;
- the outcome of the research must be published by the accrediting body or any body to which it validly sub-delegates its power. This is a separate condition to the public interest test;
- when the outcome of the research is published, the accrediting body must publish the query used by the researcher to interrogate the data; and
- the accrediting body, or any body it delegates its power to, must from time to time publish criteria by reference to which it will determine whether to grant accreditation to the research; these criteria may impose additional conditions to those set out in the primary legislation. The accrediting body is obliged to consult on these criteria.

The additional conditions that the UKSA are likely to set out in the Code of Practice and not in primary legislation will be consistent with the conditions currently imposed by the ADRN approvals panel, which include:

- be feasible, viable, ethical and have a clear potential public benefit;
- make a case for using administrative data to carry out the research;
- show that the data can only be accessed through the Network, rather than alternatives (for example Farr Institute, UK Data Service Secure Lab, or longitudinal studies); and
- not be research which a government department or agency would carry out as part of its normal operations

Annex B

Interaction with existing law and practice

The DPA and Human Rights Act will continue to apply in full to each proposed processing activity and data shares made under this Bill. No data sharing under this Bill should breach the Data Protection Act. S33 DPA has a well-established legislative provision that provides certain exemptions from the provisions of the DPA for data processed for research purposes and we would expect the provisions in the proposed Bill to build on or cross refer to this existing provision. In addition a researcher should be able to rely on the legitimising condition in Schedule 2 paragraph 6, that the “processing is necessary for the purposes of legitimate interests pursued by the data controller.”

The approach adopted is one of enabling the bodies concerned to adopt and operate this model within the existing legal framework. The following restrictions on disclosure by public authorities should not be overruled by the legislation. The reason for this is that this would not be acceptable from a privacy protection point of view:

- Human Rights Act 1998;
- Data Protection Act 1998 – where the processing is of personal data and the principles set out in the schedule 1 to DPA are not met;
- Intellectual property rights – where the use of the information in the manner proposed will amount to a breach;
- Official Secrets Act – where disclosure would cause an individual who is subject to the Act to be in breach of the duties contained in the Act; and
- FOIA – where data would be withheld in response to a freedom of information request

There is no proposal to alter or amend those laws nor is there any intention to make any data sharing mandatory on any party (the system is to be entirely permissive in nature).

It is assumed that under the existing law the Source data controllers will place whatever restrictions they wish on the processing of the data, from retaining complete control under a data processing contract, through to restricting onward disclosure or location of the data or the types of processing that can be conducted with the data under a data share agreement⁴.

The Bill should make any changes necessary to provide public bodies with the necessary vires and/or removing statutory bars so that they are put in a position equivalent to a private company able to share data using this model, as long as the share complies with the DPA. Though we expect the vast majority of the structure and administrative arrangements for this new model of data-sharing to be created administratively or under existing law, we are aware that a few provisions will require new provision to be made and a statutory basis.

Reporting

As soon as possible after the end of each financial year (this follows the precedent of s27 SRSA) the UKSA must produce a report on what it has done during that year and what it intends to do during the next financial year. It must lay this report before Parliament and all the devolved legislatures. This report must be published.

Appeals process

As a matter of general legal principle, fairness and Human Rights requirements, any accreditation power provided by the legislation should be subject to a right of appeal. As the body taking the accrediting decision is a recognised public body subject to judicial review (e.g. the UK Statistics Authority) then

⁴ It is worth noting that the 1st Data protection Principle requires personal data to be processed only in a manner that is 'fair' and this includes abiding by any requirements imposed when the personal data was obtained through a data share. Consequently a breach of, for example, a no onward disclosure provision from the original data share agreement would mean a breach of the DPA even if but for that term in the share agreement it would be allowed and DPA compliant.

judicial review should provide this appeal function.