

----- Forwarded message -----

[REDACTED]

Date: 11 April 2016 at 11:24

Subject: Resopnse to Consultation - Better Use of Data

To: "data-sharing@cabinetoffice.gov.uk" <data-sharing@cabinetoffice.gov.uk>

Dear Sir/Madam,

Please note, I am not answering on behalf of the [REDACTED] but in my own personal capacity. I am a qualified (non-practising) solicitor with a Master of Laws in data protection.

My comments concern **Consultation Questions 17 and 20** (the disclosure of information for research purposes Bill/proposals), and in particular the extent to which the EU's directly-applicable GDPR is taken into account with respect to these.

I note on page 8 of the Consultation Paper, at section 25, mention is made of the upcoming GDPR, however I am concerned by the lack of reference to this in the Bill/proposals. Fundamental/Human Rights are inherent in the GDPR, and in its overall structure. Such rights are widely interpreted by the ECtHR/ECJ, with narrow exceptions permitted for any interference, and are reflected in the language of the GDPR (Council of the EU revised text dated 6 April 2016), as set out below:

Article 9(2)(g)

processing is necessary for reasons of substantial public interest, on the basis of Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject;
or

Article 9(2)(i)

processing is necessary for archiving purposes in the public interest, or scientific and historical research purposes or statistical purposes in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.

Article 89(1)

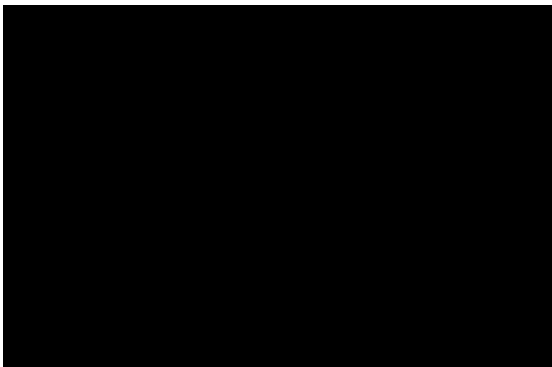
Processing of personal data for archiving purposes in the public interest, or scientific and historical research purposes or statistical purposes, shall be subject to in accordance with this Regulation appropriate safeguards for the rights and freedoms of the data subject.

All examples above require any derogations from the prohibition on the processing of special categories of personal data to provide for, and be subject to, safeguards for the rights and freedoms of the data subject. Beyond a brief mention of the ECHR Article 8 right to privacy on page 8 (section 26) of the consultation and in Annex B of the Evidence Base, I cannot see this reflected adequately in the Bill.

The GDPR contains other specific provisions relevant to the sharing of data for research, for example on Privacy by Design and Default (Article 25), Joint Data Controllers (Article 26) and Security of Processing (Article 32) – specifically in relation to pseudonymisation/encryption of personal data, which need to be taken into account, along with a definition of pseudonymisation.

Though the GDPR has achieved tripartite agreement at EU level and will become directly applicable in 2018 across all member states, the UK “Brexit” referendum decision is imminent. I believe that the UK’s position in relation to the EU will, however, have negligible effect on the EU’s requirements on protection of citizen’s data: the UK will be continually required to demonstrate a high standard of data protection and protection of fundamental rights of the data subject in order to trade; EU jurisprudence is now embedded in UK case law; and the UK will still no doubt be continued to be bound by a number of EU-like laws for some time to come, including the GDPR.

Kind regards,



Better Use of Data Consultation

Response of East Sussex County Council April 2016

Contact: Stuart Russell, Strategic Performance Manager,

Consultation questions

Improving public service delivery

1. Are there any objectives that you believe should be included in this power that would *not meet* these criteria?
No, however clarity around data sharing principals, purpose, source and ongoing governance arrangements (including consent) for information at local levels need to be understood and clarified.
2. Are there any public authorities that you consider would not fit under this definition?
No, but it may be too broad. A 'function of a public nature' for the purposes of this legislation needs to be defined. The schedule listing the specified public authorities should not include non-public authorities that may be included in the legislation in some way because they are exercising functions of a public nature on behalf of a specified public authority, see below.
3. Should non-public authorities (such as private companies and charities) that fulfil a public service function to a public authority be included in the scope of the delivering public services power?
Yes, but the non-public sector body should be required to disclose information to the public sector body for which it is fulfilling a public service function and not directly to any other bodies. Therefore non-public authorities would not be included in the schedule of specified public authorities. Legislation needs to be clear about the conditions under which non-public sector bodies will be included in the scope of the public service delivery power and how, for example when they are fulfilling a public service function to a public authority under contract or under the conditions of a grant. As above 'function of a public nature' needs to be defined. Legislation must also be clear about what non-public sector bodies are and the definition of private companies, charities, not for profit organisations and social enterprises.
4. Are these the correct principles that should be set out in the Code of Practice for this power?
Yes

Providing assistance to citizens living in fuel poverty

5. Should the Government share information with non-public sector organisations as proposed for the sole purpose of providing assistance to citizens living in fuel poverty [or for any other reason]?
Yes, it should for this sole purpose.
6. Would the provision of energy bill rebates, alongside information about energy efficiency support, be appropriate forms of assistance to citizens living in fuel poverty?
Yes
7. Are there other forms of fuel poverty assistance for citizens that should be considered for inclusion in the proposed power?
Could anything be done to ensure citizens living in fuel poverty received the lowest fuel tariff automatically. The information should also be shared with local authorities so they can target promotion of take up of energy efficiency support to landlords?

Access to civil registration to improve public service delivery

8. Should a government department be able to access birth details electronically for the purpose of providing a public service, e.g. an application for child benefit?
Yes. The information should also be available to other public authorities e.g. Clinical Commissioning Groups and local authorities so that other benefits can be applied e.g. dependent children for housing and council tax benefit claims.
9. Do you think bulk registration information, such as details of all deaths, should be shared between civil registration officials and specified public authorities to ensure records are kept up to date (e.g. to prevent correspondence being sent to families of a deceased person)?
Yes.

Overall we would like to be able to understand the impact of these proposals on our local registration service and how it is funded. Specifically we would like more detail on the fees for disclosure of information. In the example given for Child Benefit would the General Register Office get a fee from the HMRC for each birth certificate checked? In effect shifting some income from local authority registration services to the General Register Office.

How does this proposal relate to the government's current Tell Us Once service, will that be superseded by this new process?

Combating fraud against the public sector through faster and simpler access to data

10. Are there other measures which could be set out in the Code of Practice covering the proposed new power to combat fraud to strengthen the safeguards around access to data by specified public authorities?
Current measures are adequate and we have no suggestions for additional measures.
11. It is proposed that the power to improve access to information by public authorities to combat fraud will be reviewed by the Minister after a defined period of time. This time will allow for pilots to be established and outcomes and benefits evaluated. How long should the Fraud gateway be operational for before it is reviewed?
In light of the time it takes to organise this type of pilot we suggest three years; one to cover the set-up and two years to enable active data matching and analysis of the effectiveness of the system.

Improving access to data to enable better management of debt owed to the public sector

12. Which organisations should Government work with to ensure fairness is paramount when making decisions about affordability for vulnerable debtors who owe multiple debts?
All public facing organisations that require some financial assessment for the provision of services e.g. Local Authorities particularly in respect of housing benefits and social care functions. Financial help bodies such as Citizen Advice and Credit Unions may be included. Information and sharing agreements plus individual consent /confidentiality could pose some issues.
13. How can Government ensure the appropriate scrutiny so pilots under the power are effectively designed and deliver against the objectives of the power?
Government should specify criteria under each of the objectives that the pilots must deliver and proper governance arrangements need to be in place which match the objectives.
14. It is proposed that the power to improve access to information by public authorities for the purpose of better managing debt owed to government will be reviewed by the Minister after a defined period of time. This time will allow for pilots to be established and outcomes and benefits evaluated. How long should the debt power be operational for before it is reviewed?

Time will be needed to establish a detailed and complex project, allow for the debt to be identified, managed, reduced and reported on so that it is possible to evaluate the pilots against set criteria. This may take three years or more.

Access to data which must be linked and de-identified using defined processes for research purposes

15. Should fees be charged by public authorities for providing data for research purposes, and if so should there be a maximum fee permitted which is monitored by the UK Statistics Authority?

Yes fees should be charged. Providing data can be a complex and time consuming process. Based on the experience of FOIs the fee should reflect the amount of work and costs needed to make the data available. There should not be a maximum fee, rather public authorities must have the power to decline requests they consider will be unreasonably burdensome.

16. To ensure a consistent approach towards departments accepting or declining requests for disclosing information for research projects, should the UK Statistics Authority as the accreditation body publish details of rejected applications and the reasons for their rejection?

Yes, this would be useful, especially for identifying previous requests for information.

17. What principles or criteria do you think should be used to identify research that has the potential for public benefit, or research that will not be in the public benefit?

- The proposed research should state the issue being addressed and what the public benefit and/or who the public beneficiaries will be including general and specific populations or communities. The proposal should include how the possible benefit could be assessed and monitored, e.g. quantitative measures such as unit cost reductions, reduced demand for public services, services outputs; and/or qualitative measures such as improved service user/public satisfaction.
- The research proposal should explain how publicly available data has been explored and what further insights and benefits are anticipated from the additional data requested for disclosure.
- The proposed research should not be solely or primarily for the achievement of qualifications, we already receive numerous FOI requests from students effectively asking us to do work for them.
- The proposed research must include plans to share the findings and published them in an open and transparent way including anonymised or de-identified data sets.
- The proposed research should not provide direct financial benefits to a limited or private interest.
- The proposed research must include a timescale for completing the research and producing the findings within a reasonable period.
- Research findings should be shared with the public authorities that provided research data in advance of publication.
- The proposed research should not damage the public interest e.g. reducing the effectiveness and competitiveness of public service markets (e.g. the care market) or aiming to maximise profit for private interests from public service provision.

Access by UK Statistics Authority to identified data for the purpose of producing official statistics and research

18. Is two years a reasonable maximum period of time for the duration of a notice for the supply of data to the UK Statistics Authority for the purposes of producing National and official statistics and statistical research?

Yes, two years may be reasonable in most cases, but there will need to be scope for exceptions as it will depend on the circumstances required to produce the information and then the process for regular production and update of data.

19. If your business has provided a survey return to the ONS in the past we would welcome your views on:

(a) the administration burden experienced and the costs incurred in completing the survey, and

ways in which the UK Statistics Authority should seek to use the new powers to further reduce the administrative burdens on businesses who provide data to the ONS for the purposes of producing National and other official statistics.

N/A

20. What principles and factors should be considered in preparing the Code of Practice [on matters to be considered before making changes to processes that collect, store, organise or retrieve data]?

The Code of Practice must be consistent with guidance from Government Departments on how local authorities must collect, store, organise, retrieve and submit data to Government Departments for any data that is to be used in the production of national or other statistics. If this is done local authorities should not need to notify the Statistics Authority when making changes in line with guidance from Government Departments on data submitted to those departments.

C. Allowing use of data for research and for official statistics

i) Access to identified data which must be linked and de-identified using defined processes for research purposes

94. Access to more varied and better quality data can provide an improved evidence base for researchers both within and outside government, leading to better informed research and more accurate statistics for 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 and businesses face. A better understanding of how people live their lives, the patterns of need and use of different services, and the resultant outcomes would support the design and delivery of better services.
95. At present researchers are often frustrated in their efforts to access public sector-held data for research projects that have a potential public benefit. Current legislation causes uncertainty in public authorities 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. In the meantime, the possible economic or social benefit from a proposed research project could be lost.
96. In addition to and not affecting existing data sharing arrangements, new legislation is proposed to improve conditions for research in two major respects:
 - a. public authorities will have much greater clarity about what data is permitted to be shared; and
 - b. information linked under the proposed power will be subject to 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.
97. The 2012 report of the Administrative Data Taskforce (ADT) recommended the use of an established model to allow for two or more datasets to be linked for research purposes in a way that minimises the risk of an individual or organisation being identified. A number of established models are already used by public authorities. These proposals do not intend to replace or amend existing models and arrangements in place between public authorities, but sets out that data accessed under the proposed power must be linked using the ADT's recommended model. This recommended model is referred to as the 'Trusted Third Party' model. The model was explored fully during the open policy-making discussions in 2014-15, and it was agreed that access to data should be conditional on the use of such methods.
98. A Trusted Third Party model involves taking two or more different identified datasets, and removing data that could identify an individual and replacing it with a form of code or encryption. The resulting de-identified data would then be available to researchers to examine and interrogate in a secure environment. There are several variants of this model, but many of the steps required to link and de-identify data apply to all.

Example of how the Trusted Third Party model would work in practice

This example follows the fictitious case of an **accredited researcher** who is trying to explore whether there is a relationship between an economic issue and a social issue. In order to find the evidence necessary to support this project, it is necessary to link datasets relating to economic factors held by "Department A" with datasets relating to social factors held by "Department B." Both datasets include records about the same individuals, including a certain "John Smith." The Trusted Third Party model allows the researcher to confirm that the "John Smith" whose records are held by Department A is the same "John Smith" in the dataset held by Department B, without John Smith's name or any other **identifying information** (e.g. John Smith's name, date of birth and National Insurance Number) being disclosed to the researcher.

Step 1

Department A takes a copy of its dataset, removes the **identifying information** for John Smith and replaces this with a unique reference number, XYZ123. This record, together with any others from which identifying information has been removed and replaced with reference numbers, becomes the **de-identified dataset**.

Department B does the same for its dataset and tags this second **de-identified dataset** with the unique reference number ABC999 in place of John Smith's identifying information.

Step 2

Department A's **identifying information** for John Smith (e.g. John Smith's name, date of birth and National Insurance Number) is placed in a new dataset, the reference XYZ123 is added, and this dataset, the **identity dataset**, is sent to an **accredited indexer**, the 'trusted third party' who will carry out the next stage of the process.

Department B does the same with its identifying information for John Smith, adding ABC999 and sending this second identity dataset to the **accredited indexer**.

Step 3

The **accredited indexer** matches the two **identity datasets'** records, then confirms to Department A and to Department B which individual records (including those for John Smith) it has been able to match.

Step 4

Department A sends their **de-identified dataset** (from Step 1) to the **accredited access facility**, as does Department B with their **de-identified dataset**.

Step 5

The **accredited indexer** produces a list stating that ABC999 matches with XYZ123 and, having stripped out any **identifying information**, sends this list, the **index**, to the **accredited access facility**.

Step 6

The **accredited access facility**, using the **index** and the two **de-identified datasets** that it has now received from Department A and Department B (amongst which are the de-identified records relating to John Smith), now links the de-identified data together and make this available to the accredited researcher under controlled conditions.

99. Whilst these proposals have been developed to facilitate the secure linking of personal data for research purposes, it is not intended that proposals should preclude the secure processing of data that does not directly identify individuals from this power, for example property reference data. Further work is being carried out on the inclusion of such data within the scope of this power.
100. As costs may be incurred by public authorities, indexers and accredited access facilities in providing data for a research project it is proposed that the new legislation allows the researcher to be charged a fee to meet the costs. The legislation will specify that any such fees charged by public authorities are charged on a costs recovery basis only.

Question fifteen: Should fees be charged by public authorities for providing data for research purposes, and if so should there be a maximum fee permitted which is monitored by the UK Statistics Authority?

The cost recovery model will lead to uncertainties for researchers who wish to use such data.

Fixed costs of linkage may be high but then the marginal cost will be low for any future researchers using the same data. How will this be incorporated into the pricing structure? The initial one-off fixed cost should not be a barrier to linkage because the benefits in the long term far outweigh the average cost.

It would be better if there was a fixed transparent cost so that researchers can incorporate these costs into funding proposals. However, this cost needs to be not so high to discourage research.

Another model would be for any compensation for upfront costs to be met by a research body. Ultimately, any costs would be borne by such research bodies anyway through funding proposals so maybe it would be better to do this directly since this is a long term investment in a public good.

We would also like to make that point all linkages should be stored securely for future research projects. This would involve another level of data security and storage issues, as would entail the third party and the individual data providers storing paired anonymised identifiers so that matches could be reproduced easily. However, a lot of effort would be saved in the long run and would keep costs down.

Protective measures specific to this proposed power

101. Through the open policy-making process, it was agreed that identified data received either directly or indirectly by researchers, indexers and others delivering parts of the process under the power must not then be further disclosed. The legislation ensures that data from which a person's identity could be deduced is protected and that

disclosure against these conditions would be a criminal offence.

102. It was agreed through the open policy-making discussions that all researchers and bodies involved in the handling of data under the powers meet specified criteria to show they are fit and proper to fulfil their roles. Therefore, data can only be disclosed if all the participating bodies and individuals are accredited. As a result it is also necessary to designate a body to have oversight of the accreditation process.
103. Open policy-making discussions concluded that the accreditation body should accredit the indexers, the researchers, the accredited access facilities and the research itself. It is proposed that minimum accreditation requirements be set out in primary legislation. The accreditation body would develop and publish additional detailed standards and requirements for indexers, researchers and their research, and access facilities to attain and maintain this accreditation.
104. Key criteria for a suitable accreditation body were discussed as part of the open policy-making process. It was agreed that the accreditation body should have expertise in statistical research and analysis and be considered fit and proper to exercise the accreditation process. The view is that the UK Statistics Authority (Statistics Authority) is the appropriate body to carry out this function.

This needs to be coordinated across all government departments and the UK Data Service so that accreditation is not duplicated. The full process should not need to be repeated frequently but refresher courses should be provided.

105. The proposed legislation sets out that the accreditation criteria must be published by the Statistics Authority together with the grounds for revoking accreditation. Open

policy-making discussions concluded that a number of conditions were essential and needed to be set out in primary legislation. The first condition is that the individual must be fit and proper to exercise their functions. Secondly, the research question being explored must be in the public interest. Thirdly, the accreditation body must consult appropriate experts when establishing and publishing accreditation criteria for research, researchers, indexers, and access facilities. Finally, the accreditation body must also maintain and publish a register of accredited researchers, indexers, and access facilities.

Question sixteen: To ensure a consistent approach towards departments accepting or declining requests for disclosing information for research projects, should the UK Statistics Authority as the accreditation body publish details of rejected applications and the reasons for their rejection?

Yes

106. It is proposed that the Statistics Authority can delegate any part of the accreditation function to other persons or bodies, so long as they meet the requirements set out in legislation.

Power to amend the de-identification process

107. Best practices for de-identifying data are likely to change over time. The proposed legislation provides the relevant Minister with a delegated power to modify the procedures by secondary legislation. Key criteria of the de-identification process (such as the removal of the identity information before it is supplied) will be set out in primary legislation and sets out that any new procedure can only be made by regulation if it adheres to those criteria. Furthermore, the proposed legislation sets out that the Minister must consult the Statistics Authority, Information Commissioner, appropriate Ministers in the devolved administrations and other persons before making such regulations.

Criteria for determining whether research is in the public interest

108. As strongly recommended by the open policy-making group discussions, the illustrative clauses do not attempt to define "the public interest." Instead, this would be a matter for the Statistics Authority, as the designated accreditation body to determine. Public interest will be consistent with the definition of "the public good" for the production and publication of official statistics set out in the Statistics and Registration Service Act 2007 (which determines the governance and functions of the Statistics Authority). The Act describes serving the public good as:

- (a) informing the public about social and economic matters; and
- (b) assisting in the development and evaluation of public policy.

Question seventeen: What principles or criteria do you think should be used to identify research that has the potential for public benefit, or research that will not be in the public benefit?

In an ideal world, the accreditation body would not be the ones to decide whether the proposed research question is good enough. It will be very difficult for such a process to be designed in fair and transparent way because it's very difficult to measure or evaluate "public benefit". Accreditation should be about whether the researcher is fit, proper and trustworthy.

However, we recognize that some kind of criteria may be necessary to evaluate the benefit of the research. We propose that a letter of support from an independent party such as a government department or a charity would be a good criterion for identifying research that has the potential for public benefit.

Any evaluation should also allow for speculative research that is typically harder to justify is in the public benefit because the outcome is much harder to predict before the researcher has even seen the data. Such speculative research has the potential to be ground-breaking in the future and this should be recognized.

Better use of data in government consultation



Name: Professor Alison Park

Position: Director

Organisation name: CLOSER, UCL Institute of Education

Address: UCL Institute of Education, 20 Bedford Way, London, WC1H 0AL

Email: [REDACTED]

Telephone: [REDACTED]

Would you like us to treat your response as confidential?

No

Is this a personal response or an official response on behalf of your organisation?

Official response

Type of responding organisation

University or other higher education institution

Nation

England

How did you find out about this consultation?

Other:

The CLOSER consortium and other partners have been involved in discussions and the various workshops since the consultation launch announcement.

May we contact you for further information?

Yes

CLOSER
UCL Institute of Education
20 Bedford Way
London



CLOSER is funded by:

Cabinet Office consultation on 'Better use of data'

Response from the CLOSER Consortium

CLOSER (Cohorts and Longitudinal Studies Enhancement Resources) is funded by the Economic and Social Research Council (ESRC) and Medical Research Council (MRC) to maximise the use, value and impact of the UK's longitudinal studies. The network brings together eight leading longitudinal studies, the British Library and the UK Data Archive.

CLOSER's response to the consultation predominantly reflects our interest in being able to link individual administrative records to longitudinal survey data collected from that same individual. The linking of these two forms of data offers potential that exceeds the sum of the two approaches in isolation. While administrative records can provide enviable coverage and frequency, they rarely provide the richness and depth that comes from survey data collected to tackle specific research questions; conversely, survey data collection can be costly and challenging to administer, leading to missing data and possible biases which can be overcome by using administrative records to fill in the missing information. Consequently, combining the two can yield considerable benefits.

For example, recent analysis of the National Pupil Database (NPD) and Higher Education Statistics Agency (HESA) found higher participation rates in higher education among a range of groups including ethnic-minorities.¹ However, those records could not help explain these differences in participation; exploring this required linking individual administrative records to survey data collected as part of the Avon Longitudinal Study of Parents and Children and the Longitudinal Survey of Young People in England. This linkage yielded important insights about the interplay between gender, ethnicity and socioeconomic status and wider social, cultural, personal and economic factors.² It is notable that the recent House of Lords Select Committee report identifies exploring this very issue – patterns of entry to higher education among young people – as an area where “existing data is unreliable and inconsistent” and strongly supports further sharing of administrative data.³

1 Crawford C and Greaves E. (2015), “Socio-economic, ethnic and gender differences in HE participation”. *BIS Research Paper No.186*. BIS: UK, London. Available from: <https://www.gov.uk/government/collections/bis-research-paper>

2 Bowes L et al. (2015), “Understanding progression into higher education for disadvantaged and under-represented groups”. *BIS Research Paper No.229*. BIS: UK, London. Available from: <https://www.gov.uk/government/collections/bis-research-paper>

3 House of Lords Select Committee on Social Mobility (2015), ‘Overlooked and left behind: improving the transition from school to work for the majority of young people’, The Stationery Office: UK, London. Available from: <http://www.publications.parliament.uk/pa/ld201516/ldselect/ldsocmob/120/120.pdf>

Currently, the UK longitudinal study community are keen to embrace the potential offered by survey and administrative data linkage but are being hindered by disparate and opaque data access processes and procedures. Our current experiences are described very accurately in paragraph 95 of the consultation. For example, one of CLOSER's projects seeks to link survey data to records held by HMRC (only in cases survey participants had consented). Progress on this has been significantly delayed because of practical difficulties within HMRC in extracting the relevant data from their systems at a feasible cost.

We welcome and support the Cabinet Office initiative and its recognition of the difficulties that researchers face when trying to access administrative data for research of public benefit. Our comments focus on the areas within the consultation which relate to improving access to de-identified administrative records for research purposes (Section C.i.). In summary, we have seven areas of comment on the proposed legislation:

1. It is unclear whether access to administrative records in order to link to survey data would be covered within the legislation; we would strongly urge that such linkage requests are explicitly described as being within its scope.
2. In order to robustly utilise shared records it is important that these are documented to a standard that allows researchers to understand the information and the mechanisms by which the information are collated.
3. We would like to see stronger mechanisms in place to incentivise data sharing by government departments.
4. We believe that greater flexibility should be introduced as regards the mechanisms by which data linkage will be enacted. We would like to see experts from the research community involved in drawing up associated Codes of Practice.
5. We note the new considerable role for the UK Statistics Authority and urge that this is sufficiently resourced in order not to create new bottle-necks in the research process.
6. We support the decision not to define 'public interest' within the legislation.
7. The proposals should clarify whether and how researchers might access health and social care data and use these data in conjunction with survey data, potentially alongside other administrative records.

The following paragraphs expand on each of these points in turn, while also responding to three of the specific consultation questions (15-17).

- 1.

The status of survey and administrative record linkage

1.1. As outlined earlier, linking survey and administrative data permits research of considerable public benefit which offers more than either form of data can contribute alone. Currently, attempts to access the necessary administrative data encounter precisely the same problems that are articulated in the consultation. However, survey data linkage is not mentioned in the consultation. We feel this is a missed opportunity and would strongly advocate its explicit inclusion in the legislation. We would be happy to provide further input into discussions on this area.

1.2. The proposed legislation would be in line with the type of data protection processes already adopted by individual surveys in order to protect the identity and privacy of people during and after the research process.

2. Documenting shared data

2.1. The proposed legislation details the sharing of information in the form of a dataset of records and provides a mechanism by which these data may be shared. We note that in order to effectively utilise these records in a scientifically defensible manner requires researchers to receive sufficient documentation (or 'metadata') to understand the data; the way in which the data were collected, processed and the potential errors and biases that may be present and the way in which all of these factors vary across time (i.e. through the implementation of version control). We suggest that sufficient emphasis is placed on this to encourage data providers to collate robust documentation and to feel empowered to share the documentation along with the underlying records. The way administrative data has been linked also needs to be fully documented and involve best practice.

3. Incentivising data sharing

3.1. We welcome the proposal that government departments should share and link more data through appropriate mechanisms. However, we note that the legislation will be permissive in nature, rather than requiring departments to share. We would like to see consideration of stronger incentives to encourage departments to participate. In our experience, there are a number of barriers to successful data linkage, only one of which (legislative uncertainty) will be currently resolved by these proposals. Other barriers include resource (including insufficient staff with the necessary data management skills and high staff turnover) and out of date data management structures or procedures.

3.2.

The Cabinet Office is consulting on the proposal that departments charge a fee to the applicant, on a cost-recovery basis, in return for supplying data for linkage (paragraph 100, consultation question 15). We are not opposed to proportionate charges, so long as these are modest and their calculation is transparent. However, we are not convinced that all departments are set up to operate on a cost-recovery basis or, crucially, that this approach will incentivise the cultural change required within departments (our views about current barriers to data sharing are set out in the previous paragraph).

3.3. We are concerned that high fees may form a barrier to some public interest research and suggest that the UK Statistics Authority consult with the public and charity organisations who fund research about appropriate fees.

3.4. We note that, if fees are charged, a mechanism will need to be developed whereby firm quotations can be provided to researchers for grant applications to Research Councils and other bodies, with a commitment to these being honoured if the grant application is approved.

4. The data linkage mechanism

4.1. Paragraphs 97 and 98 outline the recommended model for data linkage, the Trusted Third Party model. Paragraph 107 recognises that best practices for de-identifying data are likely to change over time, and proposes that the procedures can be modified by secondary legislation. We believe that scope for revision of the exact mechanisms by which data linkage is enacted is vital, in order to avoid building in obsolescence and a disincentive to improve methodologies. We would argue that the legislation should enshrine the principle of 'effectively anonymised' rather than specifying the exact mechanism by which this should occur. We would like to see research experts and other stakeholders involved in developing a relevant Code of Practice.

4.2. Paragraph 97 cites the Administrative Data Taskforce (ADT) report and its recommendation of the Trusted Third Party model. We note that this report did not only provide one data linkage solution as it recognised that other approaches solutions might be required in particular circumstances. We also note that this report advocates data destruction after initial use. We would strongly argue against this approach, particularly if linkage between administrative and survey data becomes within the scope of the legislation. As recognised by the Information Commissioners Office's Code of Practice on anonymisation, certain research (including longitudinal research) requires retention of administrative records for an indefinite period.⁴

5.

⁴ As outlined in Section 9 of the code of practice: <https://ico.org.uk/media/for-organisations/documents/1061/anonymisation-code.pdf>

The role of the UK Statistics Authority

- 5.1. Paragraphs 104, 105 and 108 describe key activities to be carried out by the UK Statistics Authority. These include the accreditation of all active roles within the data sharing process (indexer, access facilitator and researcher), accreditation of the research (including decision-making about whether the proposed research will be in the public interest). We welcome the role of an oversight body but would urge that it is adequately resourced for this function, in order to avoid it becoming a bottleneck.
- 5.2. We note the existence of existing relevant documentation and standards which we would draw to the consultation's attention. In addition to the Information Commissioner's Office code of practice on anonymization (which we touch on in para 4.2), these include Data Safe Haven concepts⁵, ESRC guidelines regarding the access of Approved Accredited Researchers to secure data⁶, and the ISO27001 Information Security international standard.
- 5.3. It would be useful to have clarity as to whether Section 33 of the Data Protection Act (which includes some exemptions to the Act for research purposes) is compatible with any new data sharing legislation.⁷
- 5.4. If linkage to survey data is to be covered by the legislation, we would like to see it being possible for individual studies to operate as the accredited access facilitator (paragraph 202). These studies would be responsible for arranging the secure combination of participants' study and administrative records, and onward sharing these with accredited researchers to conduct accredited research.
- 5.5. The Cabinet Office is consulting on whether the UK Statistics Authority should publish details of rejected applications and the reasons for their rejection (consultation question 16). We support the principle of publishing details of why specific applications were rejected, for transparency reasons, but thought will need to be given as to how much detail about the rejected application should be made public as an application may contain valuable Intellectual Property in addition to confidential information about applicants' information security arrangements.
- 5.6.

⁵ Burton et al (2015), 'Data Safe Havens in health research and healthcare', *Bioinformatics*, 31:20.

⁶ <https://www.ukdataservice.ac.uk/media/311608/CD140-SecureAccessAgreement.pdf>

⁷ <http://www.legislation.gov.uk/ukpga/1998/29/section/33>

As outlined earlier, we have concerns about the extent to which departments will be incentivised to provide data once an application has been approved. Given this, we would urge the Cabinet Office and the UK Statistics Authority to give thought to other information that might be made publically available, and which might act as an incentive for departmental participation. Examples might include the number of applications made to specific departments for data, the number of datasets provided, time taken to do so, and so on.

6. Defining public interest

6.1. Paragraph 108 states that the legislation does not attempt to define 'the public interest'. Instead, the decision as to whether a research application meets public benefit criteria will be a task for the UK Statistics Authority. The consultation asks for views on what principles or criteria should be used to identify research that has the potential for public benefit (consultation question 17).

6.2. We support the proposal that public benefit is not defined in the legislation. Our main recommendation is that a wide definition of potential public benefit is used, and that public benefit is not linked explicitly to the improvement of public services. Research using administrative records can be of wider public benefit – for example, work on understanding the causes and consequences of social change, or exploring the impact of policy and practice on individuals and their families.

6.3. We would recommend that organisations seeking access to administrative data are required to publish outcomes from this research and place them in the public domain. We note that this commitment to publication forms part of the MRC definition of Bona Fide research.⁸

6.4. Non-public bodies engaging in research (for instance, commercial research companies and third sector organisations) are not clearly covered by the proposals. We would welcome a clear statement about whether such organisations would have access to data. We note recent Wellcome Trust research, prepared to inform the Caldicott Review, which shows considerable public support for access to health records by academic researchers, charities and organisations working with the public sector. It also identifies a significant minority which objects to commercial organisations having access to health records under any conditions.⁹ We note that concern about commercial access to health care records was a major concern in the public reaction to Care.data. Given this, we see it as vital that specific mention is made of the access arrangements and safeguards deployed if different non-public bodies were able to access administrative records.

7.

⁸ See: <http://www.nshd.mrc.ac.uk/data/data-sharing/meta-data-repository/bona-fide-research/>

⁹ Ipsos Mori (2016), 'The One-Way Mirror: Public attitudes to commercial access to health data'. Available at: www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Personal-information/Public-engagement/index.htm

Integration of health and social care data

7.1. It is not clear whether and how it will be possible to link administrative health and social care records to other records. We appreciate that the scope of this consultation in this area is very limited (as outlined in paragraph 28), as these data are subject to separate legislation and are under review by Dame Fiona Caldicott. We see it as vital that any standards for research that emerge from the Cabinet Office proposals are integrated with those that are developed for patient records. Data linkage across health and other domains underpins hugely valuable research of considerable public benefit and should not be prevented by departmental differences in data access.

Responding to the consultation

Your details

To evaluate responses properly, we need to know who is responding to the consultation and in what capacity.

We will publish our evaluation of responses. Please note that we may publish all or part of your response unless you tell us (in your answer to the confidentiality question) that you want us to treat your response as confidential. If you tell us you wish your response to be treated as confidential, we will not include your details in any published list of respondents, although we may quote from your response anonymously.

Name (optional):

Position (optional):

Organisation name:

Media, Communication and Cultural Studies Association (MeCCSA)

<http://www.meccsa.org.uk>

(MeCCSA is the subject association for academic staff and students in UK Higher Education in the fields in its title)

Address:

c/o Prof. P. Golding

Hon. Sec. MeCCSA,

4, Granville Road,

Newcastle upon Tyne

NE2 1TP

Email: [REDACTED]

Telephone (optional):

Would you like us to treat your response as confidential?*

If you answer yes, we will not include your details in any list of people or organisations that responded to the consultation.

Yes No

Is this a personal response or an official response on behalf of your organisation?

Personal response

Official response

If you ticked “Official response”, please respond accordingly:

Type of responding organisation*

Business

Charity

Local authority

Central government

Wider public sector (e.g. health bodies, schools and emergency services)

University or other higher education institution

Other representative or interest group (please answer the question below)

Type of representative group or interest group

Union

Employer or business representative group

Subject association or learned society (MeCCSA is the subject association for academic staff and students in UK Higher Education in the fields in its title)

Equality organisation or group

School, college or teacher representative group

Other (please state below)

Nation*

England

Wales

Northern Ireland

Scotland

Other EU country: _____

Non-EU country: _____

How did you find out about this consultation?

Gov.uk website

Internet search

Other

Discussion at Association Executive Committee _____

May we contact you for further information?

Yes No

Questions

We have focussed our response on the two final sections, questions 15-20, which are of direct concern to MeCCSA members.

The consultation document states that the “These proposals are not about selling public or personal data, collecting new data from citizens or weakening the Data Protection Act 1998” (p. 3). Whilst we welcome this stance, we remain concerned that once such a system of cross-referenced and linked data is established it opens the possibility of this position changing in future. In other words, safeguards must be put in place to prevent *future* commercialisation or undue and prejudicial protection of public or personal data, and erosion of the Data Protection Act, by a new Government.

The proposed “shift towards querying datasets through APIs in place of the typical practice of using bulk data shares” (p. 3) is welcome, provided there is flexibility within this to enable bulk data download should it be required for research purposes.

MeCCSA is concerned that the proposals determine that “data can only be disclosed if all the participating bodies and individuals are accredited” (point 102, p. 29) and alarmed that this extends to “the researchers, the accredited access facilities and the research itself” (point 103, p. 29). The language of the Impact Assessment pertaining to disclosure of de-identified data is vague and is already aligned towards specific disciplines in the way it describes even the most generic accreditation assumptions. A great deal of both quantitative and qualitative data on public behaviour, attitudes, or requirements in areas related to communications, culture, and media is held by statutory bodies (and it is not clear if quasi public bodies such as ofcom would be included in these proposals) and it would be enormously damaging to UK research to place any unnecessary restraint on the availability of such data.

Whilst we agree it is important that the accreditation body “should have expertise in statistical research and analysis” (point 104, p. 29), any organisation tasked with managing or preparing data should also give due consideration to qualitative research and non-statistical methods of analysis. Such data as that held in, for example, Family Spending, even the census data or in such aggregate summaries as Social Trends, are regularly and necessarily used by members. The more fundamental point is that whatever the expertise of an accrediting body, the proposed procedure would be wholly incompatible with free academic debate, in which dispute about the valid, accurate, or proper use of primary data is itself a matter for discussion in the public domain between researchers, not a matter for pre-judgement by an accrediting body, least of all one which is close to or part of the body compiling the data. The expertise of the proposed body is not the issue we question but rather its capacity improperly to pre-decide who would have access to data and how it would be used.

Better use of data presupposes making available structured data for research, though it is important that this does not restrict access to data that remains unstructured. That is, researchers must continue to be allowed access to public documents where they can extract and structure data in ways that may have been unanticipated when the legislation or accreditation guidelines were drafted. Moreover, it is important that when data is made available it is clear through meta-data as to both what data *is included* as well as being explicit about what data *is excluded*.

Access to data which must be linked and de-identified using defined processes for research purposes

Question fifteen: Should fees be charged by public authorities for providing data for research purposes, and if so should there be a maximum fee permitted which is monitored by the UK Statistics Authority?

Yes

No

In our view linked and de-identified data should be made accessible for research purposes without the imposition of a fee. Whilst the consultation documentation states that any such fee would be “charged on a costs recovery basis only” (pp. 17 and 30), we remain concerned that this would place a prohibitive burden upon researchers within our field. There is scant detail about what a “commercial rate” for cost recovery might entail (either by the public body or the third party Indexer), and it is likely any such fee would restrict access to all but large-scale funded projects (e.g. RCUK or Horizon 2020). There is no certainty that any such costs would be accepted in research tenders by such bodies, and research inevitably and frequently results in data requirements that cannot be anticipated at the design stage. Such a system would undermine humanities research in particular, where research is often carried out by individual academics supported by institutional or QR funding allocations, and could have seriously deleterious impacts on preliminary, unfunded, or pilot research. In the very unwelcome event that a fee is nevertheless introduced, we strongly advocate the introduction of a category for non-funded research that offers either a fee-waiver or a symbolic maximum fee for providing data for research purposes.

.....
.....

.....

Question sixteen: To ensure a consistent approach towards departments accepting or declining requests for disclosing information for research projects, should the UK Statistics Authority as the accreditation body publish details of rejected applications and the reasons for their rejection?

(X) Yes

The Government should be increasing transparency of information it holds, as well as its operational procedures and decision making processes to make information that is in the “public interest” public by default. To this end we support the UKSA publishing details of rejected applications. It is important that reasons provided for rejection are clearly given and there is adequate opportunity for appeal against decisions made. We remain concerned that Government departments might use the cover of a permissive power to reject applications for research that might question current policy, or does not serve their political interests or priorities. Research ‘in the public interest’, even were this presumed to be a criterion of judgement, should not be equated to ‘aligned with current policy’.

.....

.....

.....

Question seventeen: What principles or criteria do you think should be used to identify research that has the potential for public benefit, or research that will not be in the public benefit?

The contribution of research to the public good is now included as a criterion of judgement in the assessment of excellence in university research undertaken by the Research Excellence Framework. It cannot and should not be pre-judged prior to the conduct of the research, and there is the very real risk that research by ‘accredited’ researchers, using data approved and released by a parti-pris accrediting body, would impede the credibility and authority of any research arising, and in turn the authority and reputation internationally of UK research. Research is either excellent or less so in its contribution to knowledge and understanding, and advance in research can only be impeded by prior selection of proposed research by holders of relevant data or their proxies. No accrediting body of the kind proposed can possibly have a credible and legitimate role in pre-judging the value of future research.

.....

.....

.....

Question twenty: What principles and factors should be considered in preparing the Code of Practice on matters to be considered before making changes to the processes that collect, store, organise or retrieve data?

- Sensitivity towards personally identifiable information
- Stressing the importance of access to data - and, crucially, metadata about what the data covers AND does not cover. The latter is all too often missed, allowing Government to swamp open data repositories with useless data, giving the impression of transparency when actual meaningful data is withheld.
- We support direct access to data for academics, to circumvent possibility of politically motivated vetoes

Responding to the consultation

Your details

To evaluate responses properly, we need to know who is responding to the consultation and in what capacity.

We will publish our evaluation of responses. Please note that we may publish all or part of your response unless you tell us (in your answer to the confidentiality question) that you want us to treat your response as confidential. If you tell us you wish your response to be treated as confidential, we will not include your details in any published list of respondents, although we may quote from your response anonymously.

Name (optional): Sharon Witherspoon

Position (optional): Acting Head of Policy

Organisation name: Academy of Social Sciences

Address: 33 Finsbury Square

London

EC2A 1AG

Email:



Telephone (optional):



Would you like us to treat your response as confidential?*

If you answer yes, we will not include your details in any list of people or organisations that responded to the consultation.

Yes No

Is this a personal response or an official response on behalf of your organisation?

Personal response

Official response

If you ticked “Official response”, please respond accordingly:

Type of responding organisation*

Business

Charity

Local authority

Central government

Wider public sector (e.g. health bodies, schools and emergency services)

University or other higher education institution

Other representative or interest group (please answer the question below)

___ Academy of Social Sciences, a professional body representing 90,000 social scientists in the UK _____

Type of representative group or interest group

Union

Employer or business representative group

Subject association or learned society

- Equality organisation or group
- School, college or teacher representative group
- Other (please state below)

___Academy of Social Sciences, representing over 90,000 social scientists in the UK _____

Nation*

- England
- Wales
- Northern Ireland
- Scotland
- Other EU country: _____
- Non-EU country: _____

How did you find out about this consultation?

- Gov.uk website
- Internet search
- Other

___Have been involved in 'open policy' deliberations convened by Cabinet Office on this issue _____

May we contact you for further information?

- Yes No

Questions

Improving public service delivery

Question one: Are there any objectives that you believe should be included in this power that would not meet these criteria?

No

Yes

If yes, please explain your reasons.

.....

.....

.....

Question two: Are there any public authorities that you consider would not fit under this definition?

No

Yes

If yes, please explain your reasons:

.....

.....

.....

Question three: Should non-public authorities (such as private companies and charities) that fulfil a public service function to a public authority be included in the scope of the delivering public services power?

Strongly agree

Agree

Neither agree nor disagree

Disagree

Strongly disagree

Please explain your reasons:

.....
.....
.....

Question four: Are these the correct principles that should be set out in the Code of Practice for this power?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

.....
.....
.....

Providing assistance to citizens living in fuel poverty

Question five: Should the government share information with non-public sector organisations as proposed for the sole purpose of providing assistance to citizens living in fuel poverty?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

.....

.....
.....

Question six: Would the provision of energy bill rebates, alongside information about energy efficiency support, be appropriate forms of assistance to citizens living in fuel poverty?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

.....
.....
.....

Question seven: Are there other forms of fuel poverty assistance that should be considered for inclusion in the proposed power?

- Yes
- No

If yes, please explain your reasons:

.....
.....
.....

Access to civil registration information to improve public service delivery

Question eight: Should a government department be able to access birth details electronically for the purpose of providing a public service, e.g. an application for child benefit?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

.....

.....

.....

Question nine: Do you think bulk registration information, such as details of all deaths, should be shared between civil registration officials and specified public authorities to ensure records are kept up to date (e.g. to prevent correspondence being sent to a deceased person)?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

.....

.....

.....

Combating fraud against the public sector through faster and simpler access to data

Question ten: Are there other measures which could be set out in the Code of Practice covering the proposed new power to combat fraud to strengthen the safeguards around access to data by specified public authorities?

Yes

No

Please explain your reasons:

.....
.....
.....

Question eleven: It is proposed that the power to improve access to information by public authorities to combat fraud would be reviewed by the Minister after a defined period of time. This time will allow for pilots to be established and outcomes and benefits evaluated. How long should the fraud gateway be operational for before it is reviewed?

.....
.....
.....

Improving access to data to enable better management of debt owed to the public sector

Question twelve: Which organisations should government work with to ensure fairness is paramount when making decisions about affordability for vulnerable debtor who owe multiple debts?

.....
.....
.....

Question thirteen: How can Government ensure the appropriate scrutiny so pilots under the power are effectively designed and deliver against the objectives of the power?

.....

.....
.....

Question fourteen: It is proposed that the power to improve access to information by public authorities to combat fraud would be reviewed by the Minister after a defined period of time. This time will allow for pilots to be established and outcomes and benefits evaluated. How long should the fraud gateway be operational for before it is reviewed??

.....
.....
.....

Access to data which must be linked and de-identified using defined processes for research purposes

Question fifteen: Should fees be charged by public authorities for providing data for research purposes, and if so should there be a maximum fee permitted which is monitored by the UK Statistics Authority?

Yes

No

.....We believe this is a public good and therefore a classic case for government funding. But if cost recovery is deemed necessary, we have set out relevant issues in paragraph 16 of our full reply to the consultation

.....
.....
.....

Question sixteen: To ensure a consistent approach towards departments accepting or declining requests for disclosing information for research projects, should the UK Statistics Authority as the accreditation body publish details of rejected applications and the reasons for their rejection?

Yes No

.....But subject to the arguments laid out in paragraph 10 of our full reply. We do not believe that full details are needed for those requests that are refused by an independent body as not meeting criteria of public good, of scientific rigour or of objectivity. We do believe that setting out details of requests approved by an independent panel with a summary of their aims and methods, the government departments concerned, the date on which the independent panel approved the request, and the dates on which the relevant government department approved the request and transferred data, or made a decision to refuse the request, along with their reasons for doing so, would be an important incentive to changing culture. Our full reply give more details.

.....

.....

.....

Question seventeen: What principles or criteria do you think should be used to identify research that has the potential for public benefit, or research that will not be in the public benefit?

See our full reply, paragraphs 17-18. We support the use of the definition of 'public good' from the Statistics and Registration Act 2007, and including statements of those procedures required to ensure that the judgement about whether or not any particular request satisfies the 'public benefit' test (as set out in paragraphs 4, 8, 9, and 19) of our full reply to the consultation.

.....

.....

.....

Access by UK Statistics Authority to identified data for the purpose of producing official statistics and research

Question eighteen: Is two years a reasonable maximum period of time for the duration of a notice for the supply of data to the UK Statistics Authority for the purposes of producing National and official statistics and statistical research?

Yes

No

.....

.....

.....
Question nineteen: If your business has provided a survey return to the ONS in the past we would welcome your views on:

a) the administration burden experienced and the costs incurred in completing the survey

.....
.....
.....
b) ways in which the UK Statistics Authority should seek to use the new powers to further reduce the administrative burdens on businesses who provide data to the ONS for the purposes of producing National and other official statistics

.....
.....
.....
Question twenty: What principles and factors should be considered in preparing the Code of Practice on matters to be considered before making changes to the processes that collect, store, organise or retrieve data?

..See our full reply, paragraphs 4 and especially 9. We are concerned in particular about the proposal in the Code of Practice that access not be given for 'research which a government department or agency would carry out as part of its normal operations'. The provision should focus on managerial or administrative tasks, not research that might shed light outcomes of policy, which would be captured by this clause as it is now drafted.

.....
.....
.....
Sharon Witherspoon
Acting Head of Policy
Academy of Social Sciences
22 April, 2016

Academy of Social Sciences response to Cabinet Office consultation on *Better Use of Data in government*

1. The Academy of Social Sciences is the national academy of academics, learned societies and practitioners in the social sciences. Its mission is to promote social science in the United Kingdom for public benefit. It is composed of around 1,000 individual Fellows, 41 Learned Societies, and a number of affiliate members, together representing nearly 90,000 social scientists.

Introduction

2. The Academy's focus in this response is on Section C of the Cabinet Office Consultation on *Better Use of Data*, on 'allowing use of data for research and for official statistics'. This section relates to 'de-identified' data for research and identifiable data for national statistical purposes. The Academy believes that the privacy implications in these two areas are rather different from the sharing of personally identifiable data between government departments for the other purposes set out in the consultation. At earlier stages in the open policy-making process, the Academy raised the issue of whether legislation about de-identified data for research and data for statistical purposes should be considered separately from legislation containing provisions for sharing personally identifiable data between government departments. We strongly believe this issue bears further consideration to ensure appropriate parliamentary consideration of the issues involved, as the issues in each area are complex though different from each other.
3. The Academy welcomes the prospect of legislation that may help promote clarity about the use of administrative data for research and statistical purposes. But, as we discuss below, we are not convinced that legislation alone will promote the cultural change that will, we believe, be necessary in order to reap the benefits of wider access to data for research and statistical purposes. Access to government data for research, even under existing mechanisms such as the Administrative Data Research Network, which has served as a model for some parts of the proposals contained in the consultation, has been frustratingly slow, often for very opaque reasons. We return to this issue in paragraph 10 below, in our discussion of incentives for access to data.
4. UK social science has long had a robust and transparent regime for data linkages. This involves participant consent where social surveys are being linked to administrative data, and methods of ensuring public scrutiny for linkages of administrative data under the Administrative Data Research Network (ADRN), set up by the Economic and Social Research Council and governed by a Board convened by the UK Statistics Authority. It is important to note from the outset that the ADRN has robust mechanisms for considering not only the scientific appropriateness of proposals for data linkage but also the ethical and privacy implications of any proposal, and that these involve the

participation of lay members (not just social scientists or statisticians). The ADRN publicly and transparently lists all approved projects to help enable public accountability and scrutiny. And the protections put into place involve provisions other than the nature of the de-identified data to help safeguard citizen privacy. These may be helpfully characterised as the 'five safes', following the Office for National Statistics and the ESRC-funded UK Data Service, <http://blog.ukdataservice.ac.uk/access-to-sensitive-data-for-research-the-5-safes/>. Provisions involve:

- Safe projects: for public good, as judged by a transparent mechanism independent of the researchers, and including lay members;
 - Safe data: de-identified (stripped of personal identifiers such as name and address and other data that pose a disproportionate risk of personal identification);
 - Safe people: through accredited and approved researcher schemes, involving training, clear and enforceable rules, and sanctions both of institutions and researchers who do not follow them;
 - Safe settings: secure settings for holding and processing data, following agreed international security safety standards;
 - Safe outputs: publication of results that have been specially screened to ensure that they cannot inadvertently provide information which could reveal personal identities.
5. The Academy is pleased that many of the safeguards for research and statistical use of administrative and other data proposed in the consultation are based on the models provided by UK social science, largely supported by the Economic and Social Research Council. We welcome too the clear statement of the 'key protective principles' set out in (the second) paragraph 13, on page 5 of the consultation document.

Access by the UK Statistics Authority to data for the purpose of producing official statistics and research

6. The Academy of Social Sciences supports the work of the Royal Statistical Society in promoting its Data Manifesto, which sets out the importance of robust statistics in democratic societies. We concur with them that the current framing of the draft legislation regarding access by the Office of National Statistics to government departmental data is not sufficient to ensure the access required for robust national statistics. This issue has also been brought to prominence by Sir Charles Bean's *Independent Review of UK Economic Statistics*. We agree with the recommendation in the Bean Review that ONS should have a similar right of access to that accorded to the Office of Budget Responsibility, with an appropriate form of independent oversight in contested cases. We note too that stronger data sharing rights are not only important for the production of timely and accurate national statistics, but are essential to achieve the reduction in respondent burden and the financial savings envisaged in the consultation document.

Access to de-identified data for research purposes

7. The Academy understands why access to de-identified data for research purposes should be based on a permissive power given to government departments and data owners. But we have concerns that the current reluctance on the part of many departments to provide data for research that has been through approved mechanisms such as the ADRN (about which the ESRC could provide details) may persist. If this legislation is to promote real access, and not just rhetoric, will, we believe require incentives for cultural change.
8. This is despite public views (see the report commissioned by the ESRC and ONS on the issue of using administrative data for research, <https://www.ipsos-mori.com/Assets/Docs/Publications/sri-dialogue-on-data-2014.pdf>) supporting the use of administrative and other data for research purposes as long as appropriate safeguards are in place. The Academy understands the complexity of making data available for social scientists' use that may address issues of contemporary social policy. It supports the use of independent and peer review mechanisms to ensure that proposed projects using government data (and indeed all data) are not only scientifically robust but provide objective analyses of complex issues. But we are concerned that governments of any complexion are often concerned that wider use of data may bring government policies under scrutiny, and the legislation could provide important incentives to promote cultural change.
9. This issue may be illustrated by the proposed Code of Practice for statistics and research, on page 18, in which it discusses the additional conditions that the UKSA are likely to set out for research, including that any proposals 'not be research which a government department or agency would carry out as part of its normal operations'. The Academy is concerned that this provision is too widely drawn. As currently drafted, government departments could justify decisions not to allow data access if the project were on an issue that was of any policy relevance. Given that data will always be provided in arrears, after a lengthy proposal process, and strict scrutiny of its objectivity, we think a redrafting is needed to ensure that data requests are not denied because the relevant department might normally carry out such work, or might carry out such work but might not publish it. The public benefit test should be sufficient. We believe that this clause should be re-drafted to focus on 'wholly managerial or administrative operational matters', to ensure that policy research was not included.
10. Other proposals in the Consultation Paper may help address the issue of incentives for permitting access. The first relates to listing proposals for data access. **Question 16** of the consultation asks whether the UK Statistics Authority, as the accrediting authority, should publish details of rejected applications for data access, along with the reasons for their rejection. We would not favour listing full details (for instance, including researcher names)

of those applications rejected by the independent adjudication mechanism as used by ADRN, which have fallen at the hurdles of scientific rigour or objectivity or public benefit. There could however be a full list with details of all proposals approved by an independent body (like the mechanisms set up by the ADRN), including a brief description of the proposal, showing which department's data was requested, and whether the final departmental decision was to allow access or not. It would also be helpful to have information about the date at which initial approval by the independent panel was granted, and the date when data was provided by the relevant government department or the decision was made not to do so, including the reasons given if access for an approved project was not granted. Such a report, if compiled on an annual basis, with oversight and publication by UKSA and perhaps put forward as part of its annual parliamentary report, would ensure that there was continuing scrutiny of the departmental practices and their reasoning, and that government departments faced some accountability mechanism in making their decisions about whether or not to grant access to requests that have been approved by the independent mechanism as meeting the public good test (and being scientifically robust).

11. In any event, we stress the importance of resourcing the UKSA appropriately for this new role in supporting the adjudication mechanism for this work.

Details of proposed data access mechanisms

12. We welcome the careful consideration of the mechanisms by which data requests might be made, the sort of scrutiny they will need, and the mechanisms for providing data contained in the Consultation Paper and the Impact Statement on de-identified data for research. We have some concerns that these are perhaps too specifically drawn in some cases (for instance in paragraphs 97 and 98, referring to Trusted Third Parties, the current mechanism for de-identification and data linkage). We wonder if the drafting should be reconsidered, setting out a standard at least as strong as the Trusted Third Party model, leaving details to be set out in secondary legislation (as proposed) but not putting detail which could become obsolete into the proposed Bill.
13. The Academy is particularly concerned that destruction of *de-identified* data not be mandated in legislation but left to secondary legislation. Given the administrative resources required for each access, and the fact that research uses may not always require the most recent data, we can see possibilities for re-use of de-identified data (after going through the process for independent scrutiny of each proposal for re-use and subject to the relevant security provisions for de-identified data).
14. This holds in particular if the proposed legislation brings the linkage of survey data and administrative data within its scope, as we would hope. To date, some of the richest uses of departmental data have arisen from the links

between UK's unrivalled collection of longitudinal survey data (such as those run from the ISER at the University of Essex and the birth cohort studies overseen by CLOSER, at University College London – Institute of Education) and administrative data. Once the administrative data are de-identified, and are linked to the survey data, they are held in secure settings and adhere to the other five 'safes' listed on paragraph 4 above. We do not believe that this contravenes the key principles set out by the Consultation Paper.

15. The Academy notes that the possible inclusion of health and social care data (including data held by the Health and Social Care Information Centre) is awaiting the findings of Dame Fiona Caldicott's review, but we think it important that if there are to be two regimes they should permit the linkage of administrative data for social research with health data, so that unintended obstacles are not created by the legislation. We believe that linking social administrative and survey data with health data will become increasingly important in addressing many issues in preventative health care and behavioural change, an issue of potentially enormous public benefit.
16. **Question 15** in the Consultation Paper raises the issue of fees that might be charged by public authorities to provide data for research purposes. The Academy understands why cost-recovery may be an issue. Ideally we would urge that the UKSA be given funds to support this work for approved projects, at least for the first few years. In principle we would make the point that funding research for public good is an appropriate role for government, and we suspect the total sums involved would not be large. But if fees are to be charged, there needs to be some work, with appropriate oversight, to establish a clear and appropriate (and we think in reality modest) tariff for provision of data. We think it important that any tariff reflect only the marginal costs of the data preparation work and that there is some scrutiny and cross-departmental consistency in this. Otherwise, departments could set inappropriately high charges to deter requests, including asking researchers to bear the costs of protracted departmental deliberations. Since most of the research requests that might make use of data for public benefit would be funded by research councils (funded by government) or research charities who fund such work as a public good, the Academy believes it is in the public interest, if costs are not to be subsidised by government as a public good, that they are kept to a minimum, and a clear tariff is set in advance that allows research funders to plan accordingly.

Definition of public benefit

17. Regarding **Question 17** of the Consultation, the Academy agrees that the definition of 'public good' as set out in the Statistics and Registration Act 2007, resting on 'informing the public about social and economic matters' and 'assisting in the development and evaluation of public policy', is a sound substantive approach, and encompasses a wide range of possible uses.

18. The Academy has long been a proponent of the view that the concept of 'public benefit' should be drawn to include certain procedural safeguards. These include that decisions about the public benefit test for permitting access should be defined by adherence to a process including consideration of the proposal independently of the researcher or funder, and involve lay members of the public (that is, not social scientists or government staff), who must decide if there is a legitimate public benefit, and balance the benefits and any potential harms. Further, we believe that an important safeguard is to publish the decisions about both recommendations that access be given, and subsequent departmental decisions either to permit or deny access, as set out in paragraph 10 above. This allows public scrutiny and a mechanism for accountability. A final test would be that research findings should be published, so placing them in the public domain.
19. The Academy believes that these tests might allow further discussion of allowing access to 'commercial users' doing public benefit research. We note the recent publication of a report by the Wellcome Foundation, to inform the deliberations of the Caldicott review, again carried out by Ipsos MORI, on public attitudes to commercial access to health data. (See <https://www.ipsos-mori.com/researchpublications/publications/1803/Commercial-access-to-health-data.aspx>.)
20. That report may help in due course to distinguish between research *users* and research *uses*, and raises many important issues. We believe that adjudication by an independent body with lay members, open listing of permissions given and requests granted, and publication of research results are a starting point in judging whether there is a genuine public benefit, even in the case of commercial users.
21. But public views on this matter have been affected by the reaction to Care.data. It will be important for any legislation to be clear about whether commercial users are or are not included; their inclusion at this stage may require further time and discussion of the Caldicott review. If this were to lead to a protracted delay in putting forward legislation, it would be better to restrict this legislation to approved users as set out by the ESRC and other research councils.

Academy of Social Sciences
April 2016



28 Bedford Square
London WC1B 3JS
Telephone: +44 (0)20 7631 0566
Fax: +44 (0)20 7323 4877
www.nuffieldfoundation.org
Registered Charity 206601

Better Use of Data in Government Consultation 2016
Government Digital Service
6th Floor
Aviation House
London
WC2B 6NH

By email: to data-sharing@cabinetoffice.gov.uk

22 April 2016

Dear Sir/Madam

This letter sets out the Nuffield Foundation's response to the consultation on better use of data in government.

The Nuffield Foundation is a charitable trust established by William Morris, Lord Nuffield, the founder of Morris Motors. Our aim is to improve social well-being and we do this by:

- Funding research and innovation projects in education and social policy, with a strong emphasis on quantitative analysis using existing (and government-managed) datasets. In 2015 we funded 34 new projects with a total value of £5 million. Seventy-five per cent of these projects were based in universities.
- Building research capacity in science and social science, most notably through Q-Step, a £19.5m programme designed to promote a step-change in quantitative social science training for undergraduates (co-funded with the ESRC and the Higher Education Funding Council for England (HEFCE)), and Nuffield Research Placements, which provide over 1,000 Year 12 students each year with the opportunity to undertake STEM research projects.

Relevance to our work

Many of the research projects that we fund draw upon administrative data, and researchers are often looking for novel ways to use multiple sets of data to inform policy. As such, access to, and analysis of, data which is gathered by government departments, arm's length bodies and allied third parties, is essential to our work. We also recognise that in many of the policy areas in which data can shed light or (at least) help to design future studies, there is no single source of data, and therefore there is great potential for aligning and joining datasets. The recent Nuffield-funded research on the 'human capital' of graduates is a pertinent

example of this¹. For the first time, researchers were able to use anonymised tax data and student loan records to look at how graduate earnings vary by institution of study, degree subject and parental income up to ten years after graduation. Of course, we are aware that some data are more sensitive than others and that different protocols might affect how these data are treated compared with, say, 'open data'.

Some areas of the consultation were more relevant to our work than others and for that reason we are responding to a sub-set of questions rather than the entire suite. The remainder of this letter sets out the views of the Nuffield Foundation on these substantive areas, offering views on the general (consultation questions 1-4) and the specific (consultation questions 15-20) areas.

General points (consultation questions 1-4)

It is our belief that issues relating to data access and linking should start from a presumption of permissiveness and that whichever authority is brought into being to manage access should adopt that position. Data requests should, of course, make a convincing case for the data required, but there should be clarity over the evaluation criteria and an onus on the relevant body to make a comprehensible and visible case for refusing requests.

We welcome the safeguards put in place to ensure that Ministers must consult widely before adding new objectives (although a-c on page 11 of the consultation seem so broad in scope as to make this power redundant). Ensuring this is not subject to 'whim' is an important step. We support the availability of data for research purposes, and it is also our view that under this legislation, the position of non-public sector bodies undertaking public services (for example Free Schools and Academies) needs to be consistent with that on public sector bodies, subject to controls over the security of data used by such organisations.

There appears to be no reference to the extent to which longitudinal and survey work can be included in the scope of this legislation and we would suggest that it should be. Linking such exercises to administrative data, for example, could be very powerful, and would have a number of benefits. First, use of administrative data for the purposes of sampling (as opposed to analysis) can be an efficient and cost-effective way of enabling researchers to identify populations of special interest who may be missing from general household surveys. This has the dual advantage of making better use of limited research budgets (which will be doubly important if new charges are proposed for data access) and of enabling the participation of hard-to-reach groups (whose under-representation in surveys is not always due to different response modes – some groups are missed because they are hard to find in the first place). Second, where survey and administrative data can be linked, this can improve data quality through triangulation of self-reported responses with administrative data, and ease respondent burden through reducing the number of questions asked, or the number of times a person is surveyed (if they can subsequently be monitored through administrative data).

We appreciate that these uses of administrative data for survey purposes raise slightly different challenges, because the researcher may need to access personal – as well as anonymised – data, but we think it is important that they receive greater attention. To that

¹ <http://www.ifs.org.uk/publications/8233>

end the Foundation, in partnership with the Economic and Social Research Council (ESRC) and the Medical Research Council (MRC)-funded CLOSER unit at University College London (UCL), is holding a seminar on 13 July to explore what might be done to make progress on these issues.

Equally, the consultation needs to be closely aligned with the protocols and considerations over the security of individuals' data which are emerging from the Caldicott review.² It is likely that principles which are fit for the sharing of health and social care will be transferable to other forms of personal data.

We also take the view that the legislation needs to avoid tying itself closely to forms of technology (for example, Application Programming Interfaces (API)) and specific mechanisms for linking data (such as described in paragraph 98). This is because the technology is moving so fast that there are risks that it will not stand the test of time. The legislation needs to focus upon comprehensible and forward-looking principles, with examples of what to use and how to do things made available in accompanying guidance (which can be updated as necessary).

There is no timetable proposed for implementation of the legislation and whilst that may be indicative of caution, it would be helpful to have some outline so that the implications can be planned for. It is also our view that the consultation document conveys the impression that data and intelligence collection are a burden and that the benefits are solely economic or efficiency inclined. As with the recent Office for National Statistics (ONS) consultation 'on Changes to ONS Products 2015', we feel that the broader social benefits (the 'public good') of intelligent data gathering, linking and sharing (with appropriate safeguards) need to be given stronger weighting.

Specific points (consultation questions 15-20)

Many of the above points have a direct bearing upon the ways in which researchers would like to use data gathered by government (for example, the legislation should include linking to survey data). However, the questions raised in relation to research require some additional scrutiny.

The creation of a body to oversee 'trusted third parties' is an attractive concept, especially if it is also able to broker and manage sharing arrangements (with a presumption of accessibility). There are, however, some logistical and financial considerations which may influence the effectiveness of such a body. Some reassurance - perhaps drawing on the experience of the Administrative Data Service (ADS) - about the current volume of data linking requests which are made to government departments and the speed of turnaround (and the costs involved) would be very helpful here.

Related to this, our main concerns would be the speed with which the requests could be processed, and the extent to which the new body can support ('compel') the development of new links. The majority of our research grants cover a period of 18-24 months, so processing requests swiftly and reliably is important. Although we have numerous positive

²

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/461993/Consultation_ND_Guardian_accessible.pdf

examples, we and our grant-holders also have considerable experience of slow and sometimes seemingly obstructive responses to data access requests from responsible bodies. This creates some concern about the extent to which a new layer of data administration will be able to act to ensure that public benefit (however that is defined, but preferably on a case-by-case basis) data requests are managed in a timely fashion.

More generally we have concerns about the creation of a system of approvals from which government departments and agencies across the UK themselves may be exempt (when they are wishing to access and analyse data). We regularly hear of instances where government departments have not adhered to rules under which they expect external researchers to operate. Whether or not these examples are anecdotal or have substance, the fact that they are in circulation does nothing to build a culture of consistency and mutual trust, which we would argue are essential foundations for responsible use of data.

We are moving forward with a proposal to develop a Family Justice Observatory,³ and this will draw heavily on data and surveys from similar sources as set out in the consultation concerning the Troubled Families programme. We are hopeful that the proposed legislation will be able to support this new enterprise in its mission to provide access to reliable research to inform analysis of options for specific cases and system performance data to assist in the allocation of resources and the development of new services and interventions. There is certainly strong demand across the system – which extends well beyond government departments – for improved understanding of both the system itself and of the drivers of demand. The Nuffield-funded work of Professor Karen Broadhurst at Lancaster University, who was able to use data from the Children & Family Court Advice and Support Service (Cafcass) to explore the extent to which mothers subject to care proceedings are recycled through the system, has brought long-overdue focus on the need to address the underlying problems that generate a high proportion of cases.

These matters link back to the resources available to any new body, and we note the suggestion about charging a fee for access and processing. Again, there are implications for a grant-making organisation such as the Nuffield Foundation as this fee will have to be included as part of any grant proposals. So clarity over the pricing structure needs to be available (and tested) early so that applicants can plan for these additional costs when making applications for funding. In principle, the Foundation is not wholly opposed to making a reasonable contribution to improved data access, but we are not convinced that doing this on a 'pay as you go' basis is the right approach, not least as the administrative burden for the data-owners is unlikely to address the need for a more permissive approach. More importantly – given that administrative data are often incomplete or of variable quality – we are concerned that a 'pay-as-you-go' model may in fact create disincentives for departments to address underlying structural/quality issues with the data if they can simply charge for the work needed to undertake temporary fixes. In addition, it will be difficult to plan for income received from a 'pay-as-you-go' model and this will make it even harder to plan for improvements to infrastructure. We also recognise that much of the effort and cost in creating new linked sets of data arises the first time it is generated. De-identified linked datasets should, as a matter of course, be made available to a wider audience to offset the costs of re-creating them.

³ <http://www.nuffieldfoundation.org/news/team-lancaster-university-and-alliance-useful-evidence-establish-scope-and-delivery-new-family->

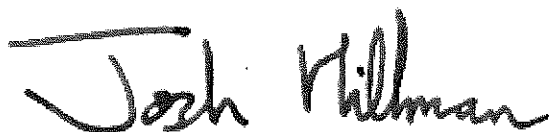
If fees are to be set, however, it would be advisable to make this scalable and to set limits based on either modelling scenarios or transferable experience (for example from Scotland). Data and linked data are obviously not a 'no cost' activity for government and some recognition of this is probably inevitable, but the introduction of fees is likely to have the effect of raising expectations of an improved service. We would therefore expect that any income generated should be directed back into service improvement and (for example) supporting the creation of generalised and regularly updated de-identified data sets (for example, the National Pupil Database – NPD).

A consequence of creating any 'accrediting' body is that it will have to police compliance with whatever code of practice it establishes and also be capable of removing recognition as an approved indexer, researcher or other data user. It is also worth considering whether this recognition could be given at organisation-level with a concomitant responsibility on that organisation to evaluate compliance with due procedures. It is not clear how frequently such re-accreditation would have to take place, but we would suggest that it should be at the very most an annual process or one which is proportionate to the track record of the organisation and/or individuals.

Lastly, the consultation asks about the publication of rejected applications. The availability of such information (without identifying the source of the request) will prove to be valuable both in terms of establishing the reliability of the application of any agreed access principles and setting out clearly what is and is not practical. Over time, this may become a reference for a working definition of 'public interest'.

We hope you will find these views useful and would be very happy to provide more details on any of the issues discussed in this response.

Yours sincerely,

A handwritten signature in black ink that reads "Josh Hillman". The signature is written in a cursive style with a horizontal line above the first few letters.

Josh Hillman
Acting Director

Responding to the consultation

Your details

To evaluate responses properly, we need to know who is responding to the consultation and in what capacity.

We will publish our evaluation of responses. Please note that we may publish all or part of your response unless you tell us (in your answer to the confidentiality question) that you want us to treat your response as confidential. If you tell us you wish your response to be treated as confidential, we will not include your details in any published list of respondents, although we may quote from your response anonymously.

Name (optional): Dr Michelle Goddard

Position (optional): Director of Policy & Standards

Organisation name: Market Research Society

Address: 15 Northburgh Street, London, EC1V 0JR

Email

[REDACTED]

Telephone (optional)

[REDACTED]

Would you like us to treat your response as confidential? No

If you answer yes, we will not include your details in any list of people or organisations that responded to the consultation.

Yes No

Is this a personal response or an official response on behalf of your organisation?

Personal response

Official response

If you ticked “Official response”, please respond accordingly:

Type of responding organisation*

Business

Charity

Local authority

Central government

Wider public sector (e.g. health bodies, schools and emergency services)

University or other higher education institution

Other representative or interest group (please answer the question below)

Type of representative group or interest group

Union

Employer or business representative group

Subject association or learned society

Equality organisation or group

School, college or teacher representative group

Other (please state below)

Nation*

England

Wales

Northern Ireland

Scotland

Other EU country: _____

Non-EU country: _____

How did you find out about this consultation?

Gov.uk website

Internet search

Other

Dods Consultation Monitoring Service

May we contact you for further information?

Yes No

Questions

Improving public service delivery

Question one: Are there any objectives that you believe should be included in this power that would not meet these criteria?

No

Yes

If yes, please explain your reasons.

.....
.....
.....

Question two: Are there any public authorities that you consider would not fit under this definition?

No

Yes

If yes, please explain your reasons:

.....
.....
.....

Question three: Should non-public authorities (such as private companies and charities) that fulfil a public service function to a public authority be included in the scope of the delivering public services power?

Strongly agree

Agree

Neither agree nor disagree

Disagree

Strongly disagree

Please explain your reasons:

.....
.....
.....

Question four: Are these the correct principles that should be set out in the Code of Practice for this power?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

.....
.....
.....

Providing assistance to citizens living in fuel poverty

Question five: Should the government share information with non-public sector organisations as proposed for the sole purpose of providing assistance to citizens living in fuel poverty?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

.....

.....
.....

Question six: Would the provision of energy bill rebates, alongside information about energy efficiency support, be appropriate forms of assistance to citizens living in fuel poverty?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

.....
.....
.....

Question seven: Are there other forms of fuel poverty assistance that should be considered for inclusion in the proposed power?

- Yes
- No

If yes, please explain your reasons:

.....
.....
.....

Access to civil registration information to improve public service delivery

Question eight: Should a government department be able to access birth details electronically for the purpose of providing a public service, e.g. an application for child benefit?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

.....

.....

.....

Question nine: Do you think bulk registration information, such as details of all deaths, should be shared between civil registration officials and specified public authorities to ensure records are kept up to date (e.g. to prevent correspondence being sent to a deceased person)?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

.....

.....

.....

Combating fraud against the public sector through faster and simpler access to data

Question ten: Are there other measures which could be set out in the Code of Practice covering the proposed new power to combat fraud to strengthen the safeguards around access to data by specified public authorities?

Yes

No

Please explain your reasons:

.....

.....

.....

Question eleven: It is proposed that the power to improve access to information by public authorities to combat fraud would be reviewed by the Minister after a defined period of time. This time will allow for pilots to be established and outcomes and benefits evaluated. How long should the fraud gateway be operational for before it is reviewed?

.....

.....

.....

Improving access to data to enable better management of debt owed to the public sector

Question twelve: Which organisations should government work with to ensure fairness is paramount when making decisions about affordability for vulnerable debtor who owe multiple debts?

.....

.....

.....

Question thirteen: How can Government ensure the appropriate scrutiny so pilots under the power are effectively designed and deliver against the objectives of the power?

.....

.....
.....

Question fourteen: It is proposed that the power to improve access to information by public authorities to combat fraud would be reviewed by the Minister after a defined period of time. This time will allow for pilots to be established and outcomes and benefits evaluated. How long should the fraud gateway be operational for before it is reviewed??

.....
.....
.....

Access to data which must be linked and de-identified using defined processes for research purposes

Question fifteen: Should fees be charged by public authorities for providing data for research purposes, and if so should there be a maximum fee permitted which is monitored by the UK Statistics Authority?

- Yes
 - No
-
.....
.....

Question sixteen: To ensure a consistent approach towards departments accepting or declining requests for disclosing information for research projects, should the UK Statistics Authority as the accreditation body publish details of rejected applications and the reasons for their rejection?

- Yes No
-
.....

.....
Question seventeen: What principles or criteria do you think should be used to identify research that has the potential for public benefit, or research that will not be in the public benefit?
.....
.....
.....

Access by UK Statistics Authority to identified data for the purpose of producing official statistics and research

Question eighteen: Is two years a reasonable maximum period of time for the duration of a notice for the supply of data to the UK Statistics Authority for the purposes of producing National and official statistics and statistical research?

Yes

No
.....
.....
.....

Question nineteen: If your business has provided a survey return to the ONS in the past we would welcome your views on:

a) the administration burden experienced and the costs incurred in completing the survey
.....
.....
.....

b) ways in which the UK Statistics Authority should seek to use the new powers to further reduce the administrative burdens on businesses who provide data to the ONS for the purposes of producing National and other official statistics
.....

.....

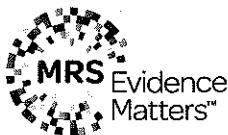
.....

Question twenty: What principles and factors should be considered in preparing the Code of Practice on matters to be considered before making changes to the processes that collect, store, organise or retrieve data?

.....

.....

.....



MRS Policy Unit

Submission to Cabinet Office Better Use of Data Consultation

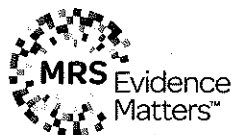
Introduction: About MRS, CGG and the research market

1. The Market Research Society (MRS) is the world's largest research association. It's for everyone with professional equity in market, social and opinion research and in business intelligence, market analysis, customer insight and consultancy. MRS supports best practice in the research industry by setting and enforcing industry standards. The Census and Geodemographics Group (CGG) is an MRS advisory board, founded in 1989 to represent the interests of this important activity. The CGG has specialists in market research, retail site location, market and database analysis, as well as census distributors and academic researchers. The CGG is involved with Census developments through representation on the ONS Business and Professional Interests Advisory Group, and with wider matters through membership of the Statistics User Forum as well as through an extensive network of contact in the market research industry.
2. The UK is the second largest research market in the world (second to the US) and the UK research sector is recognised as leading the way in the development of creative and innovative research approaches. In 2015, MRS with PWC undertook an updated assessment of the size and impact of the UK research and evidence market, *The Business of Evidence 2016*¹. One of the main findings from this report is the size of the UK 'business of evidence' market, which employs up to 73,000 people and generates £4.8 billion in annual gross value added (GVA).

About this response

3. We have taken into account the contents of the consultation paper and the questions posed and have focused our response on "Part C – Allowing use of data for research and for official statistics" which sets out proposals to allow the Office of National Statistics (ONS) to get statistical data from government and businesses and to make it easier for researchers to use data from across government.
4. We confirm that no parts of the response are confidential and that the information may be attributed to MRS.

¹ See Summary of Business of Evidence report 2016 at https://www.mrs.org.uk/pdf/boe_info.pdf.

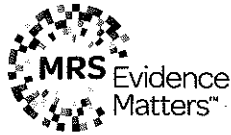


Response to Consultation Paper and Questions: Part C – Allowing use of Data for Research and Official Statistics

5. MRS and CGG welcome the move towards greater data sharing and support the key objectives of delivery of better services to citizens and better statistics so that people, organisations and government can make better decisions. Facilitating greater data sharing with appropriate ethical and legal safeguards is critical in making more effective use of existing data sets. As recognised in the 2015 Independent Review of UK Economic Statistics² it is critical for the ONS to make the most of existing and new data sources and the technologies for dealing with them.
6. Ensuring that the system has in-built safeguards to enshrine and protect the privacy of individuals will be vital in ensuring that the extension of data sharing earns the trust and confidence of citizens and assuages any concerns on the treatment and use of personal data. Against this background and contingent on the establishment of appropriate safeguards we consider that the proposals need to go further than creating a permissive regime (that does not compel data sharing) to developing a regime that compels or provides stronger incentives to encourage data sharing by public authorities.
7. In light of this the proposals for access to data need to be strengthened and we would welcome further clarity in several areas including:
 - i. *Consistency of data sharing frameworks* - As is highlighted in the consultation document the current legal framework is unclear and unwieldy with a multitude of complex and restricted legal gateways for data sharing. We consider that it is an opportune moment particularly with the implementation of new rights and responsibilities in the pending 2018 implementation of the General Data Protection Regulation (GDPR) to move towards a strengthened regime that rationalises data sharing approaches. Reform in this area should seek to simplify, rather than complicate, the pathways and demonstrate a holistic approach to data sharing in the public sector. Similarly, as is pointed

²

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/507081/2904936_Bean_Review_Web_Accessible.pdf



out in the consultation paper the development of similar approaches in Scotland, Northern Ireland and Wales would be useful in building overall consistency.

- ii. *Access by private sector researchers* - Overall it is important that these reforms create a level playing field for all researchers and equally cover researchers in both academic and private sector organisations who are engaged in the same type of analysis. We note that the paper makes no reference to private sector research/researchers and it is important that the scope is clarified so that the proposals specifically include access to data by all types of researchers across public sector, academic, charitable, private sector and media environments,
- iii. *Consistency in interpretation of key terms* – It is important that terms frequently used in a data protection context are clearly and consistently defined and used. Mechanisms should be put in place to ensure that terms such as “de-identified” and “anonymised” are used and interpreted in line with the regulatory approach of the Information Commissioner’s Office (ICO) and in line with guidance issued by ICO such as the ICO Anonymisation Code.
- iv. *Type of datasets* – The proposals are currently targeted at enabling the Office for National Statistics (ONS) to access detailed administrative data to provide more accurate, frequent and timely statistics. However survey data held across departments is also a valuable source for enriched insights and it should be clarified whether these datasets are included or intended to be included within the scope of the legislation. Additionally it is also noted that health and social care datasets are not mentioned and it would be useful to understand whether and how these data sets fit into the overall scheme.

Access to data which must be linked and de-identified using defined processes for research purposes

Question fifteen: Should fees be charged by public authorities for providing data for research purposes, and if so should there be a maximum fee permitted which is permitted by the UK Statistics Authority?

8. We agree in principle with a reasonable fee being charged by public authorities for providing data for research purposes on a costs recovery basis. Recognising that different approaches could be taken to implementation of this such as sliding scales and maximum fees the most important factor is that there is transparency.



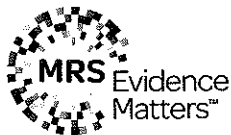
In setting the fee account must be taken of the quality of the data being provided and guidance on who bears the burden if the data is of poor quality should be considered and provided. Clarity and certainty on the amount or manner of calculation is vital in order to allow the cost of accessing research to be properly estimated and assessed in developing research proposals. Critically the same fee should apply across all sectors.

Question sixteen: To ensure a consistent approach towards departments accepting or declining requests for disclosing information for research projects, should the UK Statistics Authority as the accreditation body publish details of rejected applications and the reasons for their rejection?

9. Publication of relevant details of research projects will be important in promoting transparency of the process and should include publication of summaries of accepted applications. This may also act as a measure for assessing how well the data sharing initiative is working across different departments as the data sharing initiative will require adaptations to create a change in culture that promotes greater data sharing. Transparency and detail on applications and rejections could also encourage "cross-department competition".
10. Additionally we consider that it is reasonable for the UK Statistics Authority to publish summaries of rejected applications and the reasons for rejection. Details to be included in the summaries will need to be tailored to respect sensitivities or confidentiality and it will be important for the UK Statistics Authority to develop criteria for making these determinations. In terms of details to be published the "metrics" may be more valuable than the names particularly as naturally the reasons for the refusal may not be related to the quality of the research.

Question seventeen: What principles or criteria do you think should be used to identify research that has the potential for public benefit, or research that will not be in the public benefit?

11. Clear principles and criteria will be vital in ensuring clarity and transparency on research that has the potential for public benefit. A key point in developing appropriate and targeted approach to determination of public benefit is to ensure that it does not exclude research by private sector researchers. Public-private partnerships are increasingly important in the delivery of services and this



development should be factored into the development of the data-sharing approaches.

12. Public benefit principles should include efficiency improvement and wastage reduction. Increased efficiency is a tangible public benefit as is the development of new and innovative goods and services that serve public needs/demands. Indeed market research, is acknowledged as playing a key role in helping business to better understand consumers, customers, and to develop appropriate goods and services which are essential for economic efficiency, innovation and progress within the UK economy. Social and opinion research carried out by researchers in the private sector is widely used by government and public bodies to understand citizens' preferences and behaviours, measure impact and assist in developing appropriate policies used, for example, in improving educational, healthcare and police services.
13. As noted in our response to the ONS consultation on Approved Researchers which raised similar issues "We believe ... that in order to minimise constraints on innovation and on administrative overhead, the interpretation of 'in the public good' should be made as broad as possible consistent with the necessary constraints on the use of data. We recognise that there is a possibility of some public concern about private sector commercial users making profits from use of secure public data, and do accept that it may be necessary to ring-fence this kind of use with an appropriate requirement that the work is in the public interest. We strongly believe however that the public interest is best served by ensuring a level playing field, particularly between private companies and academic institutions. The rules should aim to avoid situations in which one sector is able to do work more easily than the other sector can do that same work."
14. We support a public benefit test that facilitates equal access by different types of researchers and consider that a policy that supports a presumption of the publication of research results could be incorporated within the framework. Publication should be subject to legal, ethical and commercial considerations and take into account reasonable grounds for withholding publication of details.



Access by UK Statistics Authority to identified data for the purpose of producing official statistics and research

15. We acknowledge the pivotal regulatory role to be carried out by the UK Statistics Authority in the oversight and the development of the accreditation scheme and accredited researchers, indexers and researchers. There is room for additional clarity in the development of this oversight role and the crafting of these principles including:

- i. Recognition of the need to be cognizant of existing guidance such as the ESRC Safe Researcher
- ii. Adherence to ethical self-regulatory professional rules should be treated as key criteria in establishing accreditation conditions. This can be achieved by incorporation of other self-regulatory frameworks such as the MRS Code of Conduct which was adopted in 1954 with the latest fully revised version of the MRS Code of Conduct coming into effect on 1 September 2014. The Code is designed to support those engaged in market research in maintaining professional standards and to reassure the general public that research is carried out in a professional and ethical manner. MRS individual members and Company Partners must comply with the Code which applies, whether they are engaged in consumer, business to business, social, opinion or any other type of research project. The commitment to uphold the MRS Code of Conduct is supported by the MRS Codeline service and a range of specialist guidelines.
- iii. Approach that facilitates continuing dialogue on the usability of the data sharing framework - Development of an advisory group or point of contact which can answer queries about what can be done with the data *once it has been received* by the requesting party is important. Sometimes ideas develop as a project evolves, and it would be useful to have people to verify uses or small changes in use from the original intention stated in the initial request. Hopefully this would also incorporate advice on whether the extension can be done under an existing agreement, or whether a new request should be submitted.



Question eighteen: Is two years a reasonable maximum period of time for the duration of a notice for the supply of data to the UK Statistics Authority for the purpose of producing National and official statistics and statistical research?

16. Establishing a maximum period of time for the duration of a notice for the supply of data is very welcome in ensuring that advance notice is given regarding changes in collection of information. We consider that two years is a reasonable maximum period.

Question nineteen: If your business has provided a survey return to the ONS in the past we would welcome your views.

17. As professional associations representing individual members we have no view on this question.

Question twenty: What principles and factors should be considered in preparing the Code of Practice?

18. The proposal to introduce legislation with applicable criminal penalties is welcome and should encourage universal compliance.

For further information or clarification on this submission please contact Dr Michelle Goddard, Director of Policy and Standards, (michelle.goddard@mrs.org.uk 020-7566-1882). This submission is made on behalf of The Market Research Society and the Census and Geodemographics Group, 15 Northburgh Street, London EC1V 0JR. The Market Research Society is a company limited by guarantee, registered in England No. 518685.

22 April 2016

Responding to the consultation

Your details

To evaluate responses properly, we need to know who is responding to the consultation and in what capacity.

We will publish our evaluation of responses. Please note that we may publish all or part of your response unless you tell us (in your answer to the confidentiality question) that you want us to treat your response as confidential. If you tell us you wish your response to be treated as confidential, we will not include your details in any published list of respondents, although we may quote from your response anonymously.

Name (optional): Professor Mark Birkin

Position (optional): Professor of Spatial Analysis and Policy

Organisation name: School of Geography

Address: University of Leeds, Leeds LS2 9JT

Email: 

Telephone (optional):

This evidence is presented on behalf of the Phase 2 Centres in the Big Data Network (BDN2) which is funded by the Economic and Social Research Council (ESRC). The network has been active since February 2014 and will operate over a five year time frame in the first instance. ESRC has invested more than £20 million of public funds in BDN2.

In relation to this consultation, the interests of BDN2 are twofold – as a user of data, many of which are generated by public agencies, for academic research; and as a secondary provider of data and associated services, to academic institutions but also to business and commercial organisations, public agencies, third sector and to the public at large. Each of the BDN2 Centres is accredited as a secure environment for safe access and use of data.

The BDN2 Network comprises three Research Centres:

- i. The Consumer Data Research Centre (CDRC) is concerned with data which are generated by retailers and other business and service organisations including (for instance) transport services, utilities and energy companies, telecommunications, and housing/ construction.

- ii. The Business and Local Government Data Research Centre (BLGDRC) collaborates with academic researchers and external partners on projects that have a positive impact on local policy, business and society.
- iii. The Urban Big Data Centre (UBDC) addresses social, economic and environmental challenges facing cities by bringing together interdisciplinary expertise of urban social scientists and data scientists to seek solutions in addressing such challenges.

BDN2 includes partners and collaborators from across the social sciences and others with an interest in economic and behavioural research e.g. in Geography, Politics, Health Sciences, Environmental Science, Transport Studies, Social Policy and Business Schools.

Phase One of the Big Data Network comprises four national centres (in England, Wales, Scotland and Northern Ireland) with a specific remit around administrative data. A separate response has been provided by this Administrative Data Research Network (ADRN).

Would you like us to treat your response as confidential?*

If you answer yes, we will not include your details in any list of people or organisations that responded to the consultation.

Yes No

Is this a personal response or an official response on behalf of your organisation?

Personal response

Official response

If you ticked “Official response”, please respond accordingly:

Type of responding organisation*

Business

Charity

Local authority

Central government

Wider public sector (e.g. health bodies, schools and emergency services)

University or other higher education institution

Other representative or interest group (please answer the question below)

Type of representative group or interest group

- Union
- Employer or business representative group
- Subject association or learned society
- Equality organisation or group
- School, college or teacher representative group
- Other (please state below)

Nation*

- England
- Wales
- Northern Ireland
- Scotland
- Other EU country: _____
- Non-EU country: _____

How did you find out about this consultation?

- Gov.uk website
- Internet search
- Other

May we contact you for further information?

- Yes No

Questions

Improving public service delivery

Question one: Are there any objectives that you believe should be included in this power that would not meet these criteria?

No

Yes

If yes, please explain your reasons.

Notwithstanding the scope of this section in relation to 'improving public service delivery' there is a danger that the powers are excessively focused on the achievement of short-term objectives (in targeting, provision and improved well-being). It is not clear that there is anything here which encourages use of 'the rich datasets available to government ... to produce world-leading research' as stated in the Ministerial Foreword.

A concrete example from our own research would be a proposal that data from the Student Loans Company could be used, possibly including linkage with tax records and Higher Education Statistics, in order to understand long-term skills migration, for example in data science¹. Such analysis might support decision-making around many problems ranging from HS2 to the funding of University Technical Colleges but would not necessarily easily satisfy the immediate public benefit test which is implied here.

Question one relates directly to paragraph 39 of the proposals which proposes 3 criteria for which data is made available. However, there is no reference at all to access to data for research purposes in this section. It could be argued that delivering benefit to individuals and improving services are aims of research and so it is covered, but it would be better if was explicitly stated, for example:

"(d) to enable access to data for the purposes of research which will enhance public authorities' ability to deliver public benefit"

Question two: Are there any public authorities that you consider would not fit under this definition?

No

Yes

¹ Government Office for Science (2016) Future of Cities: Graduate Mobility and Productivity

If yes, please explain your reasons:

The stated definition is of 'a person who exercises functions of a public nature'. Presumably 'organisation' could be substituted for 'person' without any loss of sense. The question of what is public or in the public benefit could presumably range extremely widely, for example in the construction of infrastructure, leisure services (e.g. cinemas, bars and restaurants), or even retail organisations such as supermarkets (in which lack of high quality provision is known to lead to high prices, poor choices and inferior diet). If there were any sense in which the legislation could be extended to embrace a wider variety of organisations, even as a set of voluntary principles, then this could increase its value significantly.

Question three: Should non-public authorities (such as private companies and charities) that fulfil a public service function to a public authority be included in the scope of the delivering public services power?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

The consultation is right to note that public benefit is not necessarily circumscribed by public agencies and the associated data. The problem of fuel poverty (page 14) is a neat illustration of the principle that third party data e.g. under the ownership of energy companies, is hugely valuable in both decision-making and research around public interest questions. Similar examples could easily be provided in relation to housing and land use, transportation, food deserts, active lifestyles, and indeed the provision of libraries, arts and cultural services. Thus we strongly support the suggestion that 'the scope of the power to cover non-public sector bodies' is helpful and should be pushed as far as is reasonably practicable to cover many facets of public policy.

Question four: Are these the correct principles that should be set out in the Code of Practice for this power?

- Strongly agree
- Agree

Neither agree nor disagree

Disagree

Strongly disagree

Please explain your reasons:

Agreed, but with the reservation that in this context, the meaning and intent of the expression 'business case for data sharing' (paragraph 44b) is somewhat unclear. We restate our support for the principle that the costs of data sharing should normally be recovered on a profit-neutral basis at most.

Providing assistance to citizens living in fuel poverty

Although fuel poverty is relevant to our research, we have no specific views to add in this section beyond those which are expressed elsewhere.

Question five: Should the government share information with non-public sector organisations as proposed for the sole purpose of providing assistance to citizens living in fuel poverty?

Strongly agree

Agree

Neither agree nor disagree

Disagree

Strongly disagree

Please explain your reasons:

.....

.....

.....

Question six: Would the provision of energy bill rebates, alongside information about energy efficiency support, be appropriate forms of assistance to citizens living in fuel poverty?

Strongly agree

Agree

Neither agree nor disagree

Disagree

Strongly disagree

Please explain your reasons:

.....
.....
.....

Question seven: Are there other forms of fuel poverty assistance that should be considered for inclusion in the proposed power?

Yes

No

If yes, please explain your reasons:

.....
.....
.....

Access to civil registration information to improve public service delivery

We have no specific views on the use of civil registration information beyond those which are expressed elsewhere.

Question eight: Should a government department be able to access birth details electronically for the purpose of providing a public service, e.g. an application for child benefit?

Strongly agree

Agree

Neither agree nor disagree

Disagree

Strongly disagree

Please explain your reasons:

.....

.....

.....

Question nine: Do you think bulk registration information, such as details of all deaths, should be shared between civil registration officials and specified public authorities to ensure records are kept up to date (e.g. to prevent correspondence being sent to a deceased person)?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

.....

.....

.....

Combating fraud against the public sector through faster and simpler access to data

Although the problem of fraud is of interest to our research community, we have no specific views to add in this section beyond those which are expressed elsewhere.

Question ten: Are there other measures which could be set out in the Code of Practice covering the proposed new power to combat fraud to strengthen the safeguards around access to data by specified public authorities?

- Yes
- No

Please explain your reasons:

.....

.....

.....

Question eleven: It is proposed that the power to improve access to information by public authorities to combat fraud would be reviewed by the Minister after a defined period of time. This time will allow for pilots to be established and outcomes and benefits evaluated. How long should the fraud gateway be operational for before it is reviewed?

.....

.....

.....

Improving access to data to enable better management of debt owed to the public sector

Although better management of debt is relevant to some researchers, we have no specific views to add in this section beyond those which are expressed elsewhere.

Question twelve: Which organisations should government work with to ensure fairness is paramount when making decisions about affordability for vulnerable debtor who owe multiple debts?

.....

.....

.....

Question thirteen: How can Government ensure the appropriate scrutiny so pilots under the power are effectively designed and deliver against the objectives of the power?

.....

.....

.....

Question fourteen: It is proposed that the power to improve access to information by public authorities to combat fraud would be reviewed by the Minister after a defined period of time. This time will allow for pilots to be established and outcomes and benefits evaluated. How long should the fraud gateway be operational for before it is reviewed??

.....
.....
.....

Access to data which must be linked and de-identified using defined processes for research purposes

We welcome the proposal to make public data more readily available for research users. We endorse warmly the suggestion that such data are potentially of immense interest and value to the research community.

We believe the proposals could do more to recognise the long-term nature of the academic research process and the value of curiosity-driven or speculative investigations. Public benefit is not necessarily either short-term or subject to the realisation of immediate value.

The proposals state that data will only be disclosed for a specific purpose and there should only be very limited circumstances under which such data is used more widely (para 38). There is no consideration as to how this fits with data sharing initiatives such as the ESRC BDN2 Centres.

Data of public interest are by no means limited to those held by public organisations. Any measures that can improve access to related datasets would be of considerable value. We support the recommendations of Sir Charles Bean's recently published 'Independent Review of UK Economic Statistics', in particular regarding stronger powers for ONS to access business data (paragraph 4.155). We anticipate that enhanced capability for UKSA/ ONS to access more data from commercial sources could assist our own applications to companies for data for research purposes.

Public data which are to be exploited in the public interest should generally be charged on a cost recovery basis, but recognising that additional costs may be incurred to promote widespread adoption in the public interest. ESRC's Big Data Network is a potential conduit between the owners and users of public data. Existing infrastructure investments in BDN could help to defray the costs and maximise the uptake of public data arising from the introduction of new legislative powers.

Question fifteen: Should fees be charged by public authorities for providing data for research purposes, and if so should there be a maximum fee permitted which is monitored by the UK Statistics Authority?

Yes

No

We accept the idea of paying towards the costs incurred by public bodies when supplying data for a research project. In general we endorse the principle which is stated in paragraph 100 that 'fees charged by public authorities are charged on a costs recovery basis only', and this approach is currently applied by BDN2 Centres in dealing with data partners. However suppliers might also consider the costs of setting up administration for charging, and the potential benefits of the data for research, in determining their charges. Similar fee structures might need to be considered in order to promote secondary use of public data for research, for example in creating sustainable long-term business models for the Big Data Research Centres represented here. We would also like to see a clear definition of how such fees would be calculated. Not only would this engender confidence in this system through transparency, it will also provide researchers with certainty relating to the cost of data when developing research projects.

The use of trusted and secure brokers such as the BDN2 Centres would help to promote sharing of data for public benefit without escalating costs for the data providers – by offering a single point of contact and by providing a critical mass of expertise, for example in the curation of data and management of relationships with multiple end-user clients.

While we recognise the thrust of these proposals 'to facilitate the secure *linking of personal* data for research purposes' (our italics) we would stress that data does not need to be either linked or personal in order to be of value for academic research. For example, aggregated neighbourhood-level data on housing benefits could be tremendously helpful in understanding patterns of deprivation in urban areas. Similar considerations might apply in accessing research data such as DWP or HMRC. It would be tremendously helpful if these protocols were seen as promoting sharing, regardless of whether the data are linked or simply de-identified, aggregated or otherwise rendered safe.

Question sixteen: To ensure a consistent approach towards departments accepting or declining requests for disclosing information for research projects, should the UK Statistics Authority as the accreditation body publish details of rejected applications and the reasons for their rejection?

✓ Yes () No

It would be helpful in the formulation of new research proposals to make available meta-data on the rejection (and acceptance) of previous applications. Resistance to this proposal might be allayed if the unsuccessful applicant were able to choose for his/ her identity not to be made public.

Since research is essentially about innovation then any advisory capability regarding obstacles to in principle data requests, or modification of requirements in order to meet the legitimate interests of the data provider or regulator would be beneficial.

We endorse the principles in the paper around accreditation of researchers, so that data may be used safely and securely in the public interest. It will be vital that the UK Statistics Authority approval process for researchers does not become a bottleneck in the process of accessing data for research.

Question seventeen: What principles or criteria do you think should be used to identify research that has the potential for public benefit, or research that will not be in the public benefit?

Ultimately, we feel that all publically funded research in this country is oriented towards the public benefit. A limitation of the paper as currently drafted is that it focuses too strongly on the use of data for immediate benefit, for example 'to improve public service delivery, (and) combat fraud against the public sector'. We wish to encourage data-sharing which addresses longer-term research goals.

Ideally there would be a clause in the narrative (paragraphs 8-14, pp3-4) which recognises that data drives economic and social research, that this research has value, and that restrictions to data sharing suppresses research that is of public benefit, regardless of whether those benefits are immediate or to be realised in the fullness of time. This is reflected in the definition of data (page 9) as 'Information which is collected, considered and used to help decision-making' which seems unnecessarily utilitarian and restrictive.

Access by UK Statistics Authority to identified data for the purpose of producing official statistics and research

Question eighteen: Is two years a reasonable maximum period of time for the duration of a notice for the supply of data to the UK Statistics Authority for the purposes of producing National and official statistics and statistical research?

Yes

No

Yes, as a backstop, but we hope that cultural change within government and public organisations (building on recent progress with government and open data) will encourage responses which are significantly more rapid.

Question nineteen: If your business has provided a survey return to the ONS in the past we would welcome your views on:

a) the administration burden experienced and the costs incurred in completing the survey

Not applicable.

b) ways in which the UK Statistics Authority should seek to use the new powers to further reduce the administrative burdens on businesses who provide data to the ONS for the purposes of producing National and other official statistics

Not applicable.

Question twenty: What principles and factors should be considered in preparing the Code of Practice on matters to be considered before making changes to the processes that collect, store, organise or retrieve data?

We believe that the Code of Practice should apply to all of ONS's data suppliers (both public and private) and welcome criminal sanctions, to ensure universal compliance.

Responding to the consultation

Your details

To evaluate responses properly, we need to know who is responding to the consultation and in what capacity.

We will publish our evaluation of responses. Please note that we may publish all or part of your response unless you tell us (in your answer to the confidentiality question) that you want us to treat your response as confidential. If you tell us you wish your response to be treated as confidential, we will not include your details in any published list of respondents, although we may quote from your response anonymously.

Name (optional):

Professor Jane L Hutton

Position (optional):

Professor of Statistics, Department of Statistics

Organisation name:

The University of Warwick

Address:

Department of Statistics, The University of Warwick

Coventry, CV4 7AI

Email

[REDACTED]

Telephone (optional)

[REDACTED]

Would you like us to treat your response as confidential?* No

If you answer yes, we will not include your details in any list of people or organisations that responded to the consultation.

Yes No

Is this a personal response or an official response on behalf of your organisation?

Personal response

Official response

If you ticked “Official response”, please respond accordingly:

Type of responding organisation*

Business

Charity

Local authority

Central government

Wider public sector (e.g. health bodies, schools and emergency services)

University or other higher education institution

Other representative or interest group (please answer the question below)

Type of representative group or interest group

Union

Employer or business representative group

Subject association or learned society

Equality organisation or group

School, college or teacher representative group

Other (please state below)

Nation*

England

Wales

Northern Ireland

Scotland

Other EU country: _____

Non-EU country: _____

How did you find out about this consultation?

Gov.uk website

Internet search

Other

May we contact you for further information?

Yes No

Questions

Improving public service delivery

Question one: Are there any objectives that you believe should be included in this power that would not meet these criteria?

No

Yes

If yes, please explain your reasons.

Question two: Are there any public authorities that you consider would not fit under this definition?

No

Yes

If yes, please explain your reasons:

.....

.....

.....

Question three: Should non-public authorities (such as private companies and charities) that fulfil a public service function to a public authority be included in the scope of the delivering public services power?

Strongly agree

Agree

Neither agree nor disagree

Disagree

Strongly disagree

Please explain your reasons:

As the objectives are to improve public services, it makes sense to include all authorities which deliver public services. Give the intended use of the powers, and the safeguards, the opportunity to include non-public authorities is important.

Question four: Are these the correct principles that should be set out in the Code of Practice for this power?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

Principles for use and guidelines should ensure that full consideration is given to the quality of the data and the work.

Providing assistance to citizens living in fuel poverty

Question five: Should the government share information with non-public sector organisations as proposed for the sole purpose of providing assistance to citizens living in fuel poverty?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

I believe the risks, costs and benefits balance in favour of this use. I note that the consultation document included 'or any other reason?'. Any other reason would have to be clearly specified; I presume this is why it does not appear in this document.

Question six: Would the provision of energy bill rebates, alongside information about energy efficiency support, be appropriate forms of assistance to citizens living in fuel poverty?

- Strongly agree
- Agree

Neither agree nor disagree

Disagree

Strongly disagree

Please explain your reasons:

Consequences of cold homes are costly in terms of healthy, pleasant and useful lives.

Question seven: Are there other forms of fuel poverty assistance that should be considered for inclusion in the proposed power?

Yes

No

If yes, please explain your reasons:

No comment: I am not knowledgeable in this area.

Access to civil registration information to improve public service delivery

Question eight: Should a government department be able to access birth details electronically for the purpose of providing a public service, e.g. an application for child benefit?

Strongly agree

Agree

Neither agree nor disagree

Disagree

Strongly disagree

Please explain your reasons:

I regard efficient provision of benefits as important both to recipients and to taxpayers.

Question nine: Do you think bulk registration information, such as details of all deaths, should be shared between civil registration officials and specified

public authorities to ensure records are kept up to date (e.g. to prevent correspondence being sent to a deceased person)?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

I agree, but it is not clear how substantial the benefit will be. It would allow assessment of claims that some general practices claim capitation fees for dead patients. It might reduce costs by stopping benefits being paid out and then reclaimed.

Combating fraud against the public sector through faster and simpler access to data

Question ten: Are there other measures which could be set out in the Code of Practice covering the proposed new power to combat fraud to strengthen the safeguards around access to data by specified public authorities?

- Yes
- No

Please explain your reasons:

I think the team responsible for writing the proposed measures should be congratulated for thorough and clear proposals.

Question eleven: It is proposed that the power to improve access to information by public authorities to combat fraud would be reviewed by the Minister after a defined period of time. This time will allow for pilots to be established and outcomes and benefits evaluated. How long should the fraud gateway be operational for before it is reviewed?

I think that three years is a sensible time, to allow for pilots to be designed, tested, revised, used, data collected, and follow-up to assess the costs, outcomes and benefits. Any shorter time would risk poor quality information on which to base a review.

Improving access to data to enable better management of debt owed to the public sector

Question twelve: Which organisations should government work with to ensure fairness is paramount when making decisions about affordability for vulnerable debtor who owe multiple debts?

All government departments to which debts might be owed. It might well also be sensible to work with housing associations and private landlords, as it is important that people have a home. Charities which provide debt counselling might be useful. (Confidential comment: I know that private landlords are typically regarded as greedy if not corrupt. I personally let one flat to a single parent at 75% of market rates, without a deposit, and two others at 85% to 90% of market rates.)

Question thirteen: How can Government ensure the appropriate scrutiny so pilots under the power are effectively designed and deliver against the objectives of the power?

I think funding should be provided to pay for the time required for scrutiny. It might be useful for a sample of the pilot studies to be evaluated by three or four external groups, such as university departments, charities and citizens advice bureaux. I would have the same pilot studies evaluated by the different groups, as this will provide a range of responses which can inform future, more streamlined, scrutiny.

Question fourteen: It is proposed that the power to improve access to information by public authorities to combat fraud would be reviewed by the Minister after a defined period of time. This time will allow for pilots to be established and outcomes and benefits evaluated. How long should the fraud gateway be operational for before it is reviewed??

Three years, as for the fuel poverty powers.

Access to data which must be linked and de-identified using defined processes for research purposes

Question fifteen: Should fees be charged by public authorities for providing data for research purposes, and if so should there be a maximum fee permitted which is monitored by the UK Statistics Authority?

(?) Yes

(?) No

Fees which recover costs might be sensible. However, as such fees will most often be paid by Research Councils, it might be more efficient not to circulate money around within government bodies.

Question sixteen: To ensure a consistent approach towards departments accepting or declining requests for disclosing information for research projects, should the UK Statistics Authority as the accreditation body publish details of rejected applications and the reasons for their rejection?

(x) Yes () No

Publishing some details and the reasons should be effective in improving the quality of research. I have worked in health and medicine for many years, and recently in behavioural economics. That work, and the reproducibility crisis debate in academic journals highlights the need for adequately designed studies with sufficient power, and correct analysis. The Economics and Social science Research Council (under its previous name) had a statistics committee which was disbanded because it found most grant proposals were poor quality. The statistical quality of social science research is very problematic. The ESRC is still trying to attract numerate researchers. Medical research quality has improved substantially since statistical review of grant proposals and journal articles became standard.

Question seventeen: What principles or criteria do you think should be used to identify research that has the potential for public benefit, or research that will not be in the public benefit?

The main and subsidiary aims of the research must be stated, and evidence that prior research has been properly summarised in order to inform the aims and the design of the study provided. The design must then be such that the aims can be achieved to sufficient accuracy (i.e. the study should be large enough to have decent power), with due consideration to limiting bias or missing data, and a statement of what implications might be for further research or for actions to be taken. An outline of the descriptive statistical summary and the statistical analysis should be given.

A comment in the 1985 International Statistics Institute Declaration of Professional Ethics is relevant: We 'have no special entitlement to study all phenomena'.

<https://www.isi-web.org/index.php/news-from-isi/296-declaration-professionalethics-2010uk?showall=1&limitstart=>

Access by UK Statistics Authority to identified data for the purpose of producing official statistics and research

Question eighteen: Is two years a reasonable maximum period of time for the duration of a notice for the supply of data to the UK Statistics Authority for the purposes of producing National and official statistics and statistical research?

() Yes

(?) No

I am not clear about the purpose of the maximum. If data on economic or social conditions is required, I would expect it to be required for the longer term. If two years the time-scale within which data must be provided in response to a notice of supply, I would incline towards one year.

Question nineteen: If your business has provided a survey return to the ONS in the past we would welcome your views on:

a) the administration burden experienced and the costs incurred in completing the survey

.....

b) ways in which the UK Statistics Authority should seek to use the new powers to further reduce the administrative burdens on businesses who provide data to the ONS for the purposes of producing National and other official statistics

.....

Question twenty: What principles and factors should be considered in preparing the Code of Practice on matters to be considered before making changes to the processes that collect, store, organise or retrieve data?

The Code of Practice might require approved researchers to provide copies of research reports and journals articles. This might be helpful in decisions about ongoing approved researcher status. The costs or burden has to be balanced with the benefit of demonstrating the value of data, and the competence (or otherwise) of approved researchers.

I note with pleasure that the adverse impact of consent on completeness of data and hence the effectiveness of actions has been recognised (14.c of the consultation paper). We discussed informed and uninformed consent and refusal over 15 years ago (see reference [1]). It is also important to distinguish the different content of consent when dealing with data on people within organisations (see reference [2]).

Paragraph 111 of the consultation paper mentions data on births, deaths and some NHS data. Accuracy of data requires complete data or an unbiased sample. The provision for Type 1 and Type 2 objections for NHS data has two serious consequences. The first is that the subset of data provided to researchers will almost certainly be biased (see reference [3]). These biases will typically be sufficiently severe to invalidate any conclusions, and hence render the time and money spent worthless. The second is that researchers on longer-term studies which began before 2016 who follow good statistical practice when they receive routine data updates will be breaking the law. Good practice would identify individuals who are no longer included in follow-up, and consider the impact of this loss on conclusions. However, it is almost certain that these individuals are not included because they

have signed type 1 or type 2 objections. Researchers are not permitted to have this knowledge.

The restrictions with which the Health and Social Care Information Centre is now required to deal make certain kinds of research, particularly epidemiological and mortality research, difficult and potentially unreliable. A report commissioned by the Department for Transport indicated that there is no reliable information on the risk of seizures after traumatic brain injury, so it is not possible to provide evidence-based decision on whether to allow a person to have a driving licence (see reference [4]). Cerebral palsy is the most common cause of physical disability in children, but the research on how long people with cerebral palsy live, and the support they require has almost ceased because of the administrative burden of consent. With the recent restrictions on legal aid, fewer claims for damages will be brought (a typical claim for a child was for £7 million; REF 2014 impact case), but government provision of services stills needs to be planned. If 'Better use of data' can help to address the difficulties, it will be very worthwhile.

[1] J L Hutton and R E Ashcroft, (2000) 'Some popular versions of uninformed consent' *Health Care Analysis* 8:41-52.

[2] J L Hutton, (2001) 'Are distinctive ethical principles required for cluster randomised controlled trials?' *Statistics in Medicine*, 20:473-488.

[3] J L Hutton, M Eccles and J M Grimshaw, (2009) 'Ethical issues in implementation research: a discussion of the problems in achieving informed consent', *Implementation Science*, (3:52), doi=10.1186/1748-5908-3-52

[4] Carol Hawley and JL Hutton, (2010) 'Systematic review of the probability of future seizures after initial seizure or other event creating an increased future risk', <http://www.dft.gov.uk/pgr/roadsafety/research/rsrr/theme6/report5/>

From Paul Clark
Chief Executive

DDI
Email



Data Sharing Policy Team
Floor 6
Aviation House
London
WC2B 6NH

95 Promenade
Cheltenham
GL50 1HZ

Tel: [REDACTED]
Web: www.hesa.ac.uk

22 April 2016

Dear Sir or Madam,

Re: Response to the Better Use of Data consultation

I am writing to set out the response of the Higher Education Statistics Agency (HESA) to the government's Better Use of Data consultation.

HESA is the agency responsible for collecting, processing, and disseminating data and information on all aspects of higher education across the UK. We collect data from 267 Higher Education Providers (HEPs), and distribute this to a range of customers. These include the higher education Funding Councils, the relevant government departments of each of the four countries of the UK, HEPs, other public bodies and agencies with an interest in higher education, researchers, media companies, and other private companies and individuals. We also produce Official Statistics on higher education, and are regulated in this by the UK Statistics Authority. A copy of our latest corporate strategy is available from our website: <http://www.hesa.ac.uk/strategy>.

Our response is limited to those sections of the proposals that deal with research and statistics, and we are not able to comment on the other strands. In general, we are supportive of the overall direction set out in the consultation document around data sharing for research and statistical purposes. We believe that removing legal barriers to increased data sharing will ensure that the UK's national data infrastructure can keep pace with technological and analytical developments, and will benefit individuals and society through better decision-making, better policy-making, and improvements in the provision of public services. However, this needs to be done in a way that safeguards individual privacy, as articulated in the document.

HESA is already engaged in, or has contributed to, a number of projects that seek to derive policy-making benefit from sharing data. We are working closely with BIS, HMRC, DWP, and the DfE on developing the 'all-education dataset', which could be used to track progression through the education system into the labour market. HESA data was also recently used in a ground-breaking study carried out by the Institute for Fiscal Studies, analysing the impact of course of study and institution on individual graduate outcomes. This study also drew on data from HMRC and from the Student Loans Company. There is significant future potential for developing this kind of analysis, particularly through linking between higher education sector data and HMRC data.


Another example of where better data sharing would improve the targeting of policy design and intervention in higher education relates to the Unique Property Reference Number (UPRN) data. We currently gather graduate

employment location data in the Destination of Leavers of Higher Education (DLHE) survey. We are in the process of reviewing this survey, and it may be that in future we will seek to gather more location information over a longer time-period, to help contextualise tax and salary data from HMRC. Without UPRN, this process would be complex and time-consuming, but with access to UPRN data we could produce rich statistical information about the geographical dimension of graduate destinations. This would in turn be of use to government: for example, for local enterprise development, enhancing business collaboration with universities, and providing better public information for graduates. In addition, we could also enhance the statistical service we provide to the ONS by improving the quality and frequency of student location information to support their work on improving planning by local authorities, fire, ambulance, and police services.

We therefore believe that putting in place measures that can increase data sharing across government for specific, well-defined purposes will increase the effectiveness of policy development and service design. This represents a major opportunity for the future, where the most significant public policy gains will be made from the vast new possibilities being opened up through open data and big data analytics. Improved data sharing arrangements could lead to major advances in areas such as skills policy development, widening participation, making more efficient use of equipment and resources, assessing the quality of teaching provision, improving public information on higher education, and allocating resources for research priorities.

While we are responding in a general way to the consultation, there are two specific questions on which we would like to comment:

- **Question three: should non-public sector bodies (such as private companies and charities) that fulfil a public service function to a public authority be included in the scope of the public service delivery power?** HESA would be classified as such a body under the terms of the consultation document. We are a charity and a company, but operate within a statutory framework to provide important public services. In order to maximise the benefits from increased data sharing identified in the document, we suggest that the scope should be extended to include non-public sector bodies that fulfil a public service function to a public authority.
- **Question fifteen: Should fees be charged by public authorities for providing data for research purposes, and if so should there be a maximum fee permitted which is monitored by the UK Statistics Authority?** Yes, we consider that public authorities should be able to charge fees for providing data for research purposes. HESA's strategy is to make as much of its core data as possible available as open data, subject to constraints relating to individual privacy and commercial confidentiality. As our aspiration is for much of our core data to be available as open data it therefore follows that any processing or data preparation for specific purposes beyond the open data we make available would require effort and resource, and would thus incur cost. It seems reasonable that the facility should exist for HESA to recover these costs. However, we do not believe there should be a regulated maximum fee, but rather that the extent of the fee should be able to reflect the complexity of the research data being requested, the cost involved in responding to the request, and whether the research is for academic and public benefit, or for private commercial purposes. We would like to reserve the right to determine our own pricing policy, and suggest instead that a guidance note on pricing principles could be issued and updated on a regular basis, to ensure transparency. Every effort will be made by HESA to continue to bear down on these costs, and to subsidise academic research (as we already do) with income from other sources. Nevertheless, it will be important to have available a mechanism for recovering the costs of making data available for research in the future, to ensure the financial sustainability of the organisation. HESA's small scale is such that it cannot easily absorb the costs of unfunded activity.

We hope you find this response helpful. Further information about any aspect of this response, or about HESA's activities in general, is available from Paul Clark, Chief Executive 



Paul Clark
Chief Executive

Responding to the consultation

Your details

To evaluate responses properly, we need to know who is responding to the consultation and in what capacity.

We will publish our evaluation of responses. Please note that we may publish all or part of your response unless you tell us (in your answer to the confidentiality question) that you want us to treat your response as confidential. If you tell us you wish your response to be treated as confidential, we will not include your details in any published list of respondents, although we may quote from your response anonymously.

Name (optional): Directors of the Economic and Social Science Research Council's **Administrative Data Research Network**

Position (optional):

Organisation name: **Administrative Data Research Network**

Address: adrn.ac.uk/about/contact/

Email: [REDACTED] Chair of the ADRN Directors Group)

Telephone (optional): [REDACTED] Chair of the ADRN Directors Group)

Would you like us to treat your response as confidential?*

If you answer yes, we will not include your details in any list of people or organisations that responded to the consultation.

Yes No

Is this a personal response or an official response on behalf of your organisation?

Personal response

Official response

If you ticked “Official response”, please respond accordingly:

Type of responding organisation*

- Business
 - Charity
 - Local authority
 - Central government
 - Wider public sector (e.g. health bodies, schools and emergency services)
 - University or other higher education institution
 - Other representative or interest group (please answer the question below)
-

Type of representative group or interest group

- Union
 - Employer or business representative group
 - Subject association or learned society
 - Equality organisation or group
 - School, college or teacher representative group
 - Other (please state below)
-

The Administrative Data Research Network (ADRN) is a UK wide partnership between universities and national statistics agencies. It is funded by the Economic and Social Research Council (ESRC) and has been established to facilitate accredited researchers in accessing de-identified administrative data for the purpose of social and economic research. The researchers can be government departments, academics or from the third sector. The ADRN has four Administrative Data Research Centres (ADRCs), one in each UK country:

- England: led by University of Southampton
- Northern Ireland: led by Queen’s University Belfast
- Scotland: led by University of Edinburgh

- Wales: led by Swansea University

These are supported by the Administrative Data Service, led by the University of Essex, and the ADRN Board, which reports to parliament.

Nation*

England

Wales

Northern Ireland

Scotland

Other EU country: _____

Non-EU country: _____

How did you find out about this consultation?

Gov.uk website

Internet search

Other

Direct contact

May we contact you for further information?

Yes No

Part I: General Statement

We, the directors of the ESRC's Administrative Data Research Network (ADRN), welcome the proposed legislation with its aim of normalising the culture of data sharing, with appropriate safe guards, across government and agree with, in particular, the rationale underlying proposals for "Allowing use of data for research purposes and for official statistics" (Strand C).

We believe that the public perceive data collected by public authorities as a public resource and expect its utility to be maximised through wider use and application. We therefore see great potential in research that will benefit the public through the reuse of administrative data. The ADRN has made much progress in this area, in particular setting up an infrastructure that guarantees high levels of privacy and security. However, legislation would help, allowing what is technically and legally feasible to be undertaken in a more consistent and efficient way.

We are very sensitive to the fact that research with the potential for public benefit can be subject to lengthy delay due to uncertainty around data sharing. This may result in research projects being abandoned due to uncertainty per se rather than any extant barrier to data sharing with the resultant loss of potential for social and economic benefit. We believe this legislation will help in reducing this uncertainty.

Transparency should be an underlying principle for the new legislation and we would hope that this legislative change would also require that government departments make public information on all their data assets, so that researchers and members of the public can understand the potential for research use. There should also be clarity in terms of the criteria used by a government department to make decisions about data sharing, thereby increasing the consistency in their application and transparency in the decision making process. Clarity about the availability of data and the potential timescales involved for accessing data resources are essential if research is to be promoted and projects are to be effectively planned, and we hope this could be improved by this legislation.

We would like to point out that the ADRN is a service designed for government as well as for academic researchers. We hope and expect that the ADRN will become the primary mechanism to implement the research acted on de-identified administrative data through its accredited access and linkage facilities. As such we would hope that there is **explicit** permission for all linked and unlinked datasets resulting from the various government departments sharing data to be available for use through the ADRN. We therefore welcome the removal of restrictions on information only being disclosed to the Statistical Authority for statistical purposes to the broader function of research, but would hope that legislation would not restrict the Statistical Authority from sharing these data for similar research purposes, for the public good, with the ADRN or through any other appropriate mechanism. We would

in particular hope that this pathway and its application to sharing of HMRC data might facilitate broader access to tax data for wider public benefit research purposes.

We were heartened to see that the Department of Health might now become in scope of this legislative change. We would strongly support their inclusion, as without a single approach to data sharing across all departments of government there remains the risk of significant public confusion, and a less than optimum data sharing environment which might still make linking with health data impossible for some kinds of public benefit research. Research at the biosocial frontier will increasingly demand answers to questions which can only be addressed with reference to linked health and socioeconomic information and will be hampered by a lack of clear guidance and equivalent data access standards.

We would like to ask that the legislation is explicit about the legal gateway created for all organisations covered by the bill – in particular arm's length bodies such as ONS and DVLA. We would like to request that the bill explicitly states that datasets, created under the linkage that occurs under Strand A and Strand B, always be available for research under the mechanisms and controls as described within Strand C. This will ensure maximum public utility will be gained from their linkage and minimise the work required to undertake bespoke linkages under Strand C.

On the whole we would wish to see an atmosphere across government where data sharing for research in the public good is seen as more than permissive but is actually encouraged and incentivised. Also we would wish that there would be clear accountability and transparency in the decision making around data sharing with the publishing of the reasons why particular applications are turned down or delayed.

Given that this legislation brings together three data sharing uses of reasonably different purpose, we would strongly argue that strand C "Allowing the use of data for research purposes and for official statistics" should be progressed even if it were decided that others cannot be moved forward immediately.

The ADRN remit is to work with purely de-identified administrative data for public benefit research purposes. The ADRN therefore has a neutral view on some of the questions asked in this consultation. However, whenever a new dataset might be created as part of the new provisions of the act we would hope that researchers working through the ADRN would have access to the de-identified versions of the datasets. In Part II of this response we include a fuller response from the legal researchers working within the ADRN to three particular questions asked in the review.

Questions

Improving public service delivery

Question one: Are there any objectives that you believe should be included in this power that would not meet these criteria?

No

Yes

If yes, please explain your reasons.

The ADRN remit is to work with purely de-identified administrative data for public benefit research purposes. The ADRN therefore has a neutral view on this question.

Question two: Are there any public authorities that you consider would not fit under this definition?

No

Yes

If yes, please explain your reasons:

The ADRN remit is to work with purely de-identified administrative data for public benefit research purposes. The ADRN therefore has a neutral view on this question.

Question three: Should non-public authorities (such as private companies and charities) that fulfil a public service function to a public authority be included in the scope of the delivering public services power?

Strongly agree

Agree

Neither agree nor disagree

Disagree

Strongly disagree

Please explain your reasons:

See the response of the ADRN legal research teams in Part II of this response.

Question four: Are these the correct principles that should be set out in the Code of Practice for this power?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

See the response of the ADRN legal research teams in Part II of this response.

Providing assistance to citizens living in fuel poverty

Question five: Should the government share information with non-public sector organisations as proposed for the sole purpose of providing assistance to citizens living in fuel poverty?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

The ADRN remit is to work with purely de-identified administrative data for public benefit research purposes. The ADRN therefore has a neutral view on this question.

Question six: Would the provision of energy bill rebates, alongside information about energy efficiency support, be appropriate forms of assistance to citizens living in fuel poverty?

- Strongly agree

- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

The ADRN remit is to work with purely de-identified administrative data for public benefit research purposes. The ADRN therefore has a neutral view on this question.

Question seven: Are there other forms of fuel poverty assistance that should be considered for inclusion in the proposed power?

- Yes
- No

If yes, please explain your reasons:

The ADRN remit is to work with purely de-identified administrative data for public benefit research purposes. The ADRN therefore has a neutral view on this question.

Access to civil registration information to improve public service delivery

Question eight: Should a government department be able to access birth details electronically for the purpose of providing a public service, e.g. an application for child benefit?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

The ADRN remit is to work with purely de-identified administrative data for public benefit research purposes. The ADRN therefore has a neutral view on this question.

Question nine: Do you think bulk registration information, such as details of all deaths, should be shared between civil registration officials and specified public authorities to ensure records are kept up to date (e.g. to prevent correspondence being sent to a deceased person)?

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Please explain your reasons:

The ADRN remit is to work with purely de-identified administrative data for public benefit research purposes. The ADRN therefore has a neutral view on this question.

Combating fraud against the public sector through faster and simpler access to data

Question ten: Are there other measures which could be set out in the Code of Practice covering the proposed new power to combat fraud to strengthen the safeguards around access to data by specified public authorities?

- Yes
- No

Please explain your reasons:

The ADRN remit is to work with purely de-identified administrative data for public benefit research purposes. The ADRN therefore has a neutral view on this question.

Question eleven: It is proposed that the power to improve access to information by public authorities to combat fraud would be reviewed by the Minister after a defined period of time. This time will allow for pilots to be established and outcomes and benefits evaluated. How long should the fraud gateway be operational for before it is reviewed?

The ADRN remit is to work with purely de-identified administrative data for public benefit research purposes. The ADRN therefore has a neutral view on this question.

Improving access to data to enable better management of debt owed to the public sector

Question twelve: Which organisations should government work with to ensure fairness is paramount when making decisions about affordability for vulnerable debtor who owe multiple debts?

The ADRN remit is to work with purely de-identified administrative data for public benefit research purposes. The ADRN therefore has a neutral view on this question.

Question thirteen: How can Government ensure the appropriate scrutiny so pilots under the power are effectively designed and deliver against the objectives of the power?

The ADRN remit is to work with purely de-identified administrative data for public benefit research purposes. The ADRN therefore has a neutral view on this question.

Question fourteen: It is proposed that the power to improve access to information by public authorities to combat fraud would be reviewed by the Minister after a defined period of time. This time will allow for pilots to be established and outcomes and benefits evaluated. How long should the fraud gateway be operational for before it is reviewed??

The ADRN remit is to work with purely de-identified administrative data for public benefit research purposes. The ADRN therefore has a neutral view on this question.

Access to data which must be linked and de-identified using defined processes for research purposes

Question fifteen: Should fees be charged by public authorities for providing data for research purposes, and if so should there be a maximum fee permitted which is monitored by the UK Statistics Authority?

Yes

No

For research, that is in the public good (the only type of research that the ADRN can legally support), the ADRN would argue no. Given that the research has to be in the public benefit and therefore naturally allied to the goals of the public authorities providing the data and therefore generally funded out of the same public purse, then all charging would produce unnecessary extra transactional and administrative costs, on public funds. It should be noted that compared to the general spend of public authorities, the cost of data provision is small, while the cost savings associated with a significant research insight might be high. It seems much more appropriate to view that over all bodies and across all costs and benefits the actual spend on data extraction will be cost neutral.

It should also be noted that in public engagement research the 'sale' of data is viewed by the public with some concern and therefore care would need to be taken in the sense that the charge was cost recovery rather than revenue generating. It would not be at all appropriate for public bodies to be, or appear to be, gaining significant revenue from the sale of government data.

Question sixteen: To ensure a consistent approach towards departments accepting or declining requests for disclosing information for research projects, should the UK Statistics Authority as the accreditation body publish details of rejected applications and the reasons for their rejection?

(X) Yes () No

Information on both accepted and rejected applications should be published, taking care not to publish any information that is sensitive to the applicant.

Question seventeen: What principles or criteria do you think should be used to identify research that has the potential for public benefit, or research that will not be in the public benefit?

See the response of the ADRN legal research teams in Part II of this response.

Access by UK Statistics Authority to identified data for the purpose of producing official statistics and research

Question eighteen: Is two years a reasonable maximum period of time for the duration of a notice for the supply of data to the UK Statistics Authority for the purposes of producing National and official statistics and statistical research?

(X) Yes

() No

The ADRN believes that two years is a reasonable notice period to be given if a data supplier is to make changes to the relevant processes for collecting, organising, storing or retrieving information that is supplied to UK Statistics Authority for the purposes of producing National and official statistics and statistical research. However, the maximum time period to supply data for research purposes should be six months.

Question nineteen: If your business has provided a survey return to the ONS in the past we would welcome your views on:

a) the administration burden experienced and the costs incurred in completing the survey

No comment

b) ways in which the UK Statistics Authority should seek to use the new powers to further reduce the administrative burdens on businesses who provide data to the ONS for the purposes of producing National and other official statistics

No comment

Question twenty: What principles and factors should be considered in preparing the Code of Practice on matters to be considered before making changes to the processes that collect, store, organise or retrieve data?

We would hope that there is a clear statement that the principle of increasing access to, and use of, administrative data for the public good would, along with greater cataloguing and transparency, translate into an explicit expectation that such valuable resources should be routinely utilised for research with a tangible public benefit within processes and structures that ensure the maximum safety of personal data, such as with the ADRN.



Administrative Data
Research Network

An ESRC Data
Investment

Part II: Better Use of Data Consultation Paper

Professor Graeme Laurie^{1*}, Leslie Stevens^{2}, Dr Kerina Jones^{3***}, Dr Nathan Lea^{4****}**

The Economic and Social Research Council (ESRC) invested £34 million in October 2013 to establish the UK's Administrative Data Research Network (ADRN).⁵ This established Administrative Data Research Centres (ADRC) in Scotland, Wales, England and Northern Ireland, each of which would '... facilitate linkage of routinely collected administrative data, thereby stimulating opportunities for innovative research and policymaking.'⁶ **Administrative data refer to the vast range of information collected by public authorities in the course of their routine operations. This includes delivery of services such as census taking, managing income tax payments, providing child protection, performing health services, running schools, monitoring crime, allocating benefits, mediating property owner and tenant disputes, administering council housing and so forth.** A key aim of the ADRN, and Information Governance/Legal Work Packages of the individual ADRCs (in Scotland, Wales and England) is to contribute towards facilitating and safeguarding the linkage of administrative data for research purposes by developing new understandings of what good governance of administrative data requires in the interests of society as a whole, individuals and administrative data custodians.

We welcome proposals to facilitate additional secure data sharing to enhance research opportunities using administrative data. In particular, we endorse the proposals in Section 2(C)(i) of the consultation document which sets out a model for

¹ Professor of Medical Jurisprudence, Mason Institute, University of Edinburgh School of Law.

² Research Fellow for the Administrative Data Research Centre Scotland, Mason Institute, University of Edinburgh School of Law.

³ Associate Director of ADRC-Wales and Associate Professor of Health Informatics, School of Medicine, Swansea University.

⁴ Senior Research Associate, UCL Institute of Health Informatics, The Farr Institute of Health Informatics Research.

⁵ Administrative Data Research Network, 'About Us' (2015) <<http://adrn.ac.uk/about>>.

⁶ Economic and Social Research Council, 'The Big Data Family Is Born - David Willetts MP Announces the ESRC Big Data Network' (10 October 2013) <<http://www.esrc.ac.uk/news-and-events/press-releases/28673/the-big-data-family-is-born-david-willetts-mp-announces-the-esrc-big-data-network.aspx>>.

data matching and de-identification, via a Trusted Third Party, so that the resulting anonymised data can be used for research. This sort of model is currently in operation in various settings, including the Secure Anonymised Information Linkage (SAIL) databank⁷, which provides access to anonymised data understood as operating outwith the scope of current data protection law. The model does not require content details pertaining to individuals to be shared, but only demographic details for matching and anonymisation. In this way, only the original data controller holds the full identifiable dataset and only anonymised data are made available for research. This use case may in some ways be considered more straightforward than some of the other proposals under consideration where fully identifiable datasets may be required in order to provide a better service to individual citizens. Although this model is in operation for some datasets, it is either not permissible, or not perceived to be permissible for others to be processed in this way. This new proposed legislation would support further data sharing by providing clarity that secure data matching and anonymisation for research is permitted by law.

However, we are cognisant that under the forthcoming General Data Protection Regulation ('GDPR') it remains unclear how such arrangements for de-identification will be considered and in particular whether the data made available under such arrangement may be considered 'personal data' and be subject to the full scope of data protection law. This is unclear because a) pseudonymous data are considered personal data in Recital 26 of the GDPR because b) it allows individuals to be singled out across data sets (the standard is no longer solely whether data are identifiable, but also if data allow the singling out of individuals).⁸ Therefore any proposed definition for de-identified data sets⁹ should be considered in context with the standard of de-identification/anonymisation provided for in the GDPR, so as to not create inconsistencies in practice and between UK and EU law.

Nevertheless, in our work for the ADRN, and in particular from our engagement with Scottish public authorities, we have found that while the law is often held out to be the critical barrier to data sharing, the law is often *not* the problem but rather behaviours around data use and sharing driven by *perceptions* of what the law does or does not require. This contributes to overall cultures of caution within each public authority (that are indicative of more systemic, organisationally specific issues such as a lack of resources, capacity and lack of understanding of risks or of incentives for data sharing).¹⁰ Given the underlying cultural resistance to data sharing by public authorities, we are uncertain whether the legislative

⁷ Swansea University, 'SAIL - The Secure Anonymised Information Linkage Databank' (2015) <<http://www.saildatabank.com/>>.

⁸ Leslie Stevens, 'The Proposed Data Protection Regulation and Its Potential Impact on Social Sciences Research in the UK' (2015) 1 European Data Protection Law Review 6–11.

⁹ Cabinet Office, 'Illustrative Clauses: Access to de-identified Data for Research Purposes' <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/503960/illustrative_clauses_for_access_to_de-identified_data_for_research_purposes.pdf>.

¹⁰ Graeme Laurie and Leslie Stevens, 'The Administrative Data Research Centre Scotland: A Scoping Report on the Legal & Ethical Issues Arising from Access & Linkage of Administrative Data' [2014] Edinburgh School of Law Research Paper No. 2014/35 <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2487971>.

proposals, which are to provide legal permission to reuse data, are enough to outweigh all data custodians' hesitancy to undertake any perceived risks or other costs associated with administrative data use or sharing. Additionally, it remains unclear how the new proposals will work with existing legal bases for the sharing and reuse of both administrative and health data, for instance those provided in the Health and Social Care and Care Acts which apply in England. There also remains a potential incompatibility with the common law duty of confidentiality, which seems to require guidance, education and codes of practice to resolve perhaps more urgently than additional legislation that may exacerbate uncertainty around how best the common law duty (which varies across UK jurisdictions) can be upheld whilst enabling reasonable data sharing.

Furthermore, as learned from other contexts, legal sanction does not necessarily obtain the requisite social licence to support new uses of data. In the case of care.data, although the initiative was perfectly lawful, the programme was stalled (and is still delayed) due to public outcry indicating a lack of meaningful engagement with publics.¹¹ We would caution against a focus on legislative fixes to the current culture of caution surrounding data sharing if this is not also accompanied by robust engagement with broader issues of social acceptability for new uses of data. Specifically, how will these broader interests be catered to? It is not clear in what ways the proposals do enough to facilitate engagement with the relevant publics whose views must be accounted for as new uses and recipients of administrative data are facilitated by these legislative proposals. It is understood that INVOLVE facilitated an open policy-making process and external engagement, including a wide consultation and a blog.¹² It is important that this is followed through to continue to engage meaningfully with stakeholders and publics.

We note that the proposals would include a new criminal offence for 'unlawful disclosure' of data (paragraph 18). This categorisation and understanding of data misuse and any resulting harm may need to be considerably reconsidered and expanded. We hold this view as a result of research that we conducted for the UK's Nuffield Council on Bioethics and the Wellcome Trust's Expert Advisory Group on Data Access in 2014/15. We found that a far broader range of data misuse/abuse types and resulting harms/impacts occur than is currently provided for in existing data protection legislation.¹³ Within the evidence gathered, we identified eight types of misuse/abuse of data resulting in harm/impact to individuals, organisations and society. These included:

- Fabrication/falsification of data;

¹¹ Pam Carter, Graeme T Laurie and Mary Dixon-Woods, 'The Social Licence for Research: Why Care.data Ran into Trouble' [2015] *Journal of Medical Ethics* <<http://jme.bmj.com/content/early/2015/01/23/medethics-2014-102374.abstract>>.

¹² Involve, 'Data Sharing: Updates from Civil Society Engagement with the UK Government on Data Sharing' (2016) <<http://datasharing.org.uk/latest/>>.

¹³ Graeme Laurie and others, 'A Review of Evidence Relating to Harm Resulting from Uses of Health and Biomedical Data' (Nuffield Council on Bioethics and Wellcome Trust Expert Advisory Group on Data Access 2015) <<http://nuffieldbioethics.org/project/biological-health-data/evidence-gathering/>>.

- Theft of data;
- Unauthorised disclosure or access to data;
- Non-secure disposal of data;
- Unauthorised retention of data;
- Technical security failures of data;
- Loss of data;
- Non-use of data.

The evidence demonstrated that unintentional behaviours far more frequently gave rise to data incidents and these behaviours were categorised as either 'maladministration' or 'human error'; maladministration being used as a catch-all category to indicate incidents where multiple failures or lapses in information governance were observed. At present we do not see how the proposed criminalisation of the 'unlawful disclosure' of data engages with this broader understanding of data misuse/abuse types or makes provision for the prevalence of unintentional behaviours in the incidence of data breaches in the UK's public sector.¹⁴ Put otherwise, we suggest that an offence that focuses on intentional 'disclosure' as the defining criminal event will be unlikely to capture most or many of the instances in which citizens will experience harm as a result of inadvertent mishandling of their data. While we concede the important distinction between intentional and gross negligent conduct, we would argue nonetheless that the requirements of robust citizen protection dictate that any criminal offence capture both categories of behaviour.

Finally, whilst we welcome the reference in the consultation to cyber security, in distinguishing between intentional or accidental disclosure and determining criminal proceedings it is important to recognise that data breaches may occur on the basis that cyber security has been abused, circumvented or otherwise failed. It is therefore important to recognise the need to provide meaningful guidance on how best cyber security is enacted in practice. The legal requirement for security is already well defined in other statute but a comprehensive means of guiding what this means in practice remains elusive and sporadic, where compliance with best practice guidelines such as the ISO 27000 series on information security management is neither a trivial or expediently achievable task. In short – if we are to legislate criminal proceedings for misuse with comprehensive criteria for establishing this, it is vital that data sharing and reuse is not being inadvertently set up to fail – with criminal consequences. A first principle of cyber security is that people are well equipped to understand and apply safe and secure working practice, and it is important to remember that legal recourse means that a failure has occurred, so any

¹⁴ The UK's ICO routinely identifies the following categories of data breach types in their quarterly data breach trend report, which also indicate the prevalence of unintentional behaviours perpetuating such incidents: loss or theft of paperwork; data posted or faxed to an incorrect recipient; data sent by email to an incorrect recipient; insecure webpages (including hacking); loss or theft of unencrypted device.ICO, 'Data Breach Trends' (22 December 2015) <<https://ico.org.uk/action-weve-taken/data-breach-trends/>>.

additional legislation should help make clear that failure must be avoided rather than suggest only that if there is a breach, there may be criminal consequences.

More stringent measures to deal with data misuse are also contained in the recently approved text of the GDPR. Under Article 83, there is a potential for monetary penalties of €10-20M or 24% of worldwide turnover to be levied depending on the nature of the infringement.¹⁵ Whilst proper deterrents for data misuse are welcomed, there should also be a note of caution, since it is possible that concerns about penalties will result in additional caution about data sharing to the detriment of research. Mechanisms to ensure the right balance are crucial to safeguard and to facilitate proper data use.

We now go on to consider specific aspects of the proposals as reflected in questions three, four and seventeen of the consultation.

Question three: Should non-public sector bodies (such as private companies and charities) that fulfil a public service function to a public authority be included in the scope of the public service delivery power?

With many public services now being outsourced to the private sector, it is unhelpful and potentially harmful to distinguish between public and private bodies where both are performing functions of a public nature. Such a distinction would only serve as a distraction from the more pertinent and pressing issues as to when data are to be legitimately used or shared by such an organisation and under what circumstances (limitations/protections). By attempting to distinguish between private and public bodies there is a risk of applying inconsistent legal and ethical standards despite there being no substantive difference between data processing by either type of entity (if the former is involved in the delivery of public services and thus involved in the processing of individuals' personal data). Including both private organisations and public authorities in the proposals would help to ensure that the use and sharing of individuals' data facilitated by the proposed legislation are subject to the same legal and ethical principles, just as personal data are governed uniformly under the Data Protection Act 1998 regardless of the class of data controller responsible for the processing in question. Furthermore, insofar as the sharing and processing of data in the proposals relates to 'the public interest', the Data Protection Act 1998 (Schedule 2, paragraph 5(d)) specifically provides that a public interest basis for processing personal data may apply to 'any person' including either public authorities or indeed private entities which are also undertaking tasks of a public nature.

Question four: Are these the correct principles that should be set out in the

¹⁵ Council of the European Union, 'Position of the Council at first reading with a view to the adoption of a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)' (6 April 2016) <<http://data.consilium.europa.eu/doc/document/ST-5419-2016-INIT/en/pdf>>.

Code of Practice for this power?

As indicated in our overarching comments on the proposals, there is a risk that by focusing on legislative fixes to the culture of caution around the use of public sector data it may distract from addressing other important elements of support to new data use/sharing initiatives, especially the need to seek and maintain adequate social licence. We welcome that a Code of Practice would accompany each of the proposals in order to guide decisions made on the basis of the new and broad legislative gateways, such as for data sharing to facilitate public service delivery. Given the findings of the Law Commission's 2014 report on data sharing within the public sector and the ineffectiveness of 'permissive' gateways,¹⁶ it is important to consider how 'Principles for use of the power' may be interpreted within pre-existing cultures of caution which are prone to risk-averse interpretations of the law where there is a lack of clear incentive or obligation to act. As stated previously, although the law is often characterised as the crucial barrier to data sharing, more often than not, it is perceptions of the law and specific organisational factors and cultures which impede more proportionate decision-making in this area. In this sense it may be unrealistic to expect legislative proposals to 'cure' the myriad of cultural and organisation specific issues which more fully contribute to the culture of caution around data.

Furthermore, as to 'Guidance for successful implementation' we should suggest that a successful business case for data sharing under this power should require demonstrating meaningful public engagement, substantiating how relevant publics' views are incorporated into decisions taken regarding data sharing and future use. We would further suggest that the inclusion of 'interested Civil Society Organisations and independent observers' (page 20) prior to and during data sharing projects would be unlikely to obtain in all circumstances the requisite level of social licence needed to support particularly novel forms of data sharing. We suggest further that action to secure and maintain social licence requires not merely compliance with any legal regime, but the manifest demonstration of specific ways in which a stated public interest or social objective has actually been furthered through data retention and use.

Question 17: What principles or criteria do you think should be used to identify research that has the potential for public benefit, or research that will not be in the public benefit?

Data collected by public authorities are a public resource and as suggested in the context with the proposals for research use of data, each proposed use should be considered from this standpoint. This requires consideration of what it means for research to be 'in the public interest'. In this context it is important that the 'public interest' not be construed as an "either/or" proposition between considerations of

¹⁶ The Law Commission, 'Data Sharing Between Public Bodies - A Scoping Report' (2014) 79–81 <http://lawcommission.justice.gov.uk/docs/lc351_data-sharing.pdf>.

protecting individuals' privacy and broader societal interests that may be served by the proposed data use. Rather, and as reflected in the European Data Protection Directive 95/46/EC, the purpose of data protection is twofold and lies in securing *both* the robust protection of individual privacy and in facilitating the use of data, especially uses that stand to result in wider public benefit. The latter would include research uses of administrative data that can result in positive outcomes for society (for example, better public services in health, education, transportation, enhanced understanding of factors that contribute to the health and overall wellness of individuals in society, and so on).¹⁷ Research 'in the public interest' could therefore be considered in these terms:

Scientifically sound and ethically robust research based on use, linkage and reuse of administrative data is in the public interest in promoting and improving economic growth, personal and social well-being, and maximising the interests of current and future generations of citizens in the UK.¹⁸

However the public interest in this context equally represents:

Citizens' rights of privacy will be safeguarded by robust and proportionate safeguards, in recognition of the public interests served by protecting individual privacy.

Crucially, there is no definitive public interest; it is always to be determined on a case-by-case basis in line with the values and specific context in question. The substantive measure of what is and is not to be considered research in the public interest will necessarily be different in each case and can be guided according to the remit of each public authority and their particular interactions with the public. Public agencies could, and in our view should, seek to articulate what reasonably serves as the public interest in their own context; research in said public interest should be a clear, stated and transparent part of their core mission. The parameters of such research use could be explored through robust engagement with relevant stakeholders. These interactions can help indicate the acceptable boundaries and limits for which data may be used for research in the public interest and therefore inform what the public interest means for each particular public authority as translated into specific research uses of their data. This understanding of the public interest, and how it translates into research uses of public sector data, would require transparent and public commitment to using data to serve the public interest through

¹⁷ Administrative Data Taskforce, 'The UK Administrative Data Research Network: Improving Access for Research and Policy' (ESRC, MRC and Wellcome Trust 2012) 1 <http://www.esrc.ac.uk/_images/ADT-Improving-Access-for-Research-and-Policy_tcm8-24462.pdf>; 'Benefits of the Data' (Administrative Data Research Network, 2015) <<http://adrn.ac.uk/admin-data/benefits>>.

¹⁸ As originally provided in our working paper: Laurie and Stevens (n 10) 39.

research. A visible and public commitment to using public sector data when it is in the public interest to do so would help public authorities seek the crucial social licence by assuaging individuals' concerns about the likelihood of any data initiative being used for anything other than projects that will positively benefit individuals and society.¹⁹ This might be further supported by clear statements and commitments about uses for which administrative data will *not* be put, such as uses which only seek to benefit and profit private companies.

Although the public interest cannot be defined exhaustively in this (or any) context, there are established criteria for determining matters of the public interest from European law, namely, the principle of *proportionality*. Proportionality is a cornerstone principle for decision-making and regulation that requires a balance of interests, relative to an assessment of real risks and likely benefits.²⁰ Determining whether a proposed research use of data is *proportionate* to the public interest aims sought can be used to indicate whether the research may be deemed 'in the public interest' for the purposes of reliance on the proposed legislation. The principle of proportionality is supported by a clear framework of assessment provided within context of the European Convention on Human Rights and decisions taken by the European Court of Human Rights.²¹ Adapted to the current context, proportionality points towards the following questions:

- Is there a clear and knowable public interest served by the proposed research?
- Has the necessity of identifiable data been demonstrated relative to the public interest?
- Has the relative intrusiveness of data use been minimised – is the least intrusive method being used?
- Are there safeguards against abuse and misuse of the data, including sanctions?²²

¹⁹ Daniel Cameron, Sarah Pope and Michael Clemence, 'Dialogue on Data: Exploring the Public's Views on Using Administrative Data for Research Purposes' (Ipsos MORI Social Research Institute 2014) 7, 17.

²⁰ Proportionality plays a crucial role in the good governance frameworks one of us developed in the health data context. Information Governance Working Group The Scottish Health Informatics Programme, 'SHIP Guiding Principles and Best Practices' (2010) <http://www.scotship.ac.uk/sites/default/files/Reports/Guiding_Principles_and_Best_Practices_221010.pdf>; Graeme Laurie and Nayha Sethi, 'Towards Principles-Based Approaches to Governance of Health-Related Research Using Personal Data' (2013) 4 *The European Journal of Risk Regulation* 43; Nayha Sethi and Graeme T Laurie, 'Delivering Proportionate Governance in the Era of eHealth: Making Linkage and Privacy Work Together' (2013) 13 *Medical Law International* 168.

²¹ *Handyside v UK* [1976] 1 EHRR 737; *Sunday Times v United Kingdom* (1979) 2 EHRR 245; *Sunday Times v United Kingdom* (No 2) [1991] ECHR 50; *A, B & C v Ireland* (2010) 53 EHRR 13.

²² Ursula Kilkelly, *The Right to Respect for Private and Family Life: A Guide to the Implementation of Article 8 of the European Convention on Human Rights* (Council of Europe 2003) <[http://www.echr.coe.int/LibraryDocs/DG2/HRHAND/DG2-EN-HRHAND-01\(2003\).pdf](http://www.echr.coe.int/LibraryDocs/DG2/HRHAND/DG2-EN-HRHAND-01(2003).pdf)>; Douwe Korff, 'The Standard Approach Under Articles 8-11 ECHR and Article 2 ECHR' (Data Protection Conference 2009, Brussels, 2009) <http://ec.europa.eu/justice/news/events/conference_dp_2009/presentations_speeches/KORFF_Douwe_a.pdf>.

Using the principle of proportionality to guide assessments of the public interest emphasises that such decisions should be taken not only on the basis of the risks involved, but also on the basis of the potential public benefits to accrue from *using* data and any harm that might arise out of *not using data*.²³ The realisation of these public benefits should be captured and evaluated, shared and analysed, to ensure that appeals to the public interest are not merely empty rhetoric but do actually deliver to the public.

²³ The harms arising out of a failure to share data have been explored in social policy literature and in context of acute failures in social services resulting in great harm and impact to individuals and communities. 'The Victoria Climbié Inquiry' (Department of Health and Home Office 2003) <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/273183/5730.pdf>; 'Paedophile Jailed for Raping Girl' (20 April 2006) <http://news.bbc.co.uk/1/hi/england/southern_counties/4926482.stm>; Christine Bellamy and others, 'Information-Sharing and Confidentiality in Social Policy: Regulating Multi-Agency Working' (2008) 86 Public Administration 737; "Mistakes Were Made." HMIC's Review into Allegations and Intelligence Material Concerning Jimmy Savile between 1964 and 2012' (HM Inspectorate of Constabulary 2013) <<http://www.justiceinspectrates.gov.uk/hmic/media/review-into-allegations-and-intelligence-material-concerning-jimmy-savile.pdf>>; Alexis Jay, 'Independent Inquiry into Child Sexual Exploitation in Rotherham 1997-2013' (2014) <www.rotherham.gov.uk/.../independent_inquiry_cse_in_rotherham.pdf>; Press Association, '“Catastrophic Failure” Allowed Convicted Killer to Murder on Day Release' (*The Guardian*, 23 March 2015) <<http://www.theguardian.com/uk-news/2015/mar/23/catastrophic-failure-allowed-convicted-killer-to-on-day-release>>; 'Risks Associated with Sharing / Not Sharing Information' (Centre of Excellence for Information Sharing 2015) <<http://informationsharing.co.uk/wp-content/uploads/2012/07/Risks-associated-with-sharing-information.pdf>>.

Better use of data consultation - NatCen Social Research response

April 2016

About NatCen

NatCen Social Research is a charity and the UK's leading independent social research institute. Driven by the belief that social research has the power to make life better, our research works for society by providing a rich understanding of people's lives and enabling them to have a powerful and influential role in shaping social policy decisions. Our studies encompass ground breaking qualitative projects through to Britain's most complex quantitative social surveys; we are commissioned by government to collect a number of official statistics including through the Health Survey for England, English Housing Survey and National Diet and Nutrition Survey.

Contact details

Head of Press & Public Affairs: Leigh Marshall

Address:

35 Northampton Square
London
EC1V 0AX

Our work with administrative data

We use administrative data in our research, and manage access to non-archived administrative data. We are experienced in curating, managing and extracting administrative data as well as substantive survey data, and other data utilised or generated during the survey process. We are used to working on consented survey to administrative record linkage and have experience of the issues that arise when working with administrative records.

We also see huge value in the linkage of survey research and administrative data and do it in numerous research projects. To give just one example, we often link survey research to the National Pupil Database in order to better understand how aspects of a child's life that we survey relate to their educational attainment.

Summary

Our response focusses on the elements in the consultation relating to the provision of data for research purposes. In addition to answering the specific questions set out in the consultation we also note the absence of information about data sharing with the wider research community, and the possibility of linking administrative data with survey data. We also believe there is an opportunity to further develop the UK data infrastructure, with the potential of decreasing the cost of curation in the long term by collectively improving the metadata of administrative data.

Our experience of working with administrative data suggest that the proposals in the consultation risk underestimating the costs associated with the provision of data for research purposes. We also assert that rejected requests should be published, primarily because of the benefit for other researchers interested in available data. On the issue of identifying public benefit we highlight the approach used in the health and social care sector as a possible model to follow..

Consultation questions

Access to data which must be linked and de-identified using defined processes for research purposes

Question fifteen: *Should fees be charged by public authorities for providing data for research purposes, and if so should there be a maximum fee permitted which is monitored by the UK Statistics Authority?*

NatCen welcome the Government's proposals to address the legal barriers to sharing and linking of de-identified datasets for analysis by accredited researchers in a secure environment.

We agree that the cost of data sharing should be recovered. However, a cost-recovery model of estimated data work may not effectively recoup the cost of the end-to-end process. If not funded appropriately this could place a strain on public authorities that could ultimately deter them from providing the data for research purposes. The consultation could say more about some of the existing funding allocations (such as ONS and the ADRN) could better support this process.

Our own experience with the management and sharing of administrative data suggests that there is significant permanent data infrastructure that needs to be funded in order to process and deliver requests, including those which are unsuccessful. NatCen has its own approvals process to enable the safe sharing of data with external researchers. If a request is successful, the external researcher is asked to reimburse the cost of data request, unless it is a straightforward request or one we have produced beforehand. We have a Data Release Panel staffed with statisticians, researchers and data management experts who oversee the approvals, administration and data preparation stages – although not all of these functions are applicable to public authorities in this instance, we hope it

demonstrates the administrative overheads of requests for non-archived data. Even those which are unsuccessful will require one or more of the staff listed below:

- Data Release Admin:- log requests, manage communications and written agreements, provide support to enclave users and brief on Information Security procedures
- Data Release Panel:- review requests, disclosure control/check outputs
- Data Management Team:- initial assessment into the feasibility of the data request prior to approval, in addition to the data preparation itself.
- IT Support:- set up the infrastructure required for secure storage, transfer, monitor retention and secure deletion of data.

The draft text proposes a cost recovery model for the person undertaking the data work which is a large part of the overall process of data release but caution should be exercised about the quality of the data i.e. whether it is well curated and ready for extraction. In our experience, administrative data are rarely structured, documented, or quality assured in the way that survey data collection is. Cleaning administrative records has the potential to be resource intensive but the cost should decrease over time as metadata is improved and data cleaning documented.

Data management work is difficult to cost/be prescriptive about. In any new data request an initial assessment of the feasibility would be required to understand the:-

1. Volume and complexity of the request
 - No of cases/ fields or variables/ systems or source datasets
 - Level of analyses, any special programming required (i.e. derivations)
 - Indexing or matching required
2. Quality of the data requested
 - Clean, structured datasets with good metadata
 - Additional cleaning and data/metadata modification

Although we have highlighted some of the potential unforeseen costs of data linkage we do think that there should be a maximum fee charged by public authorities. Data linkage for research purposes presents a huge opportunity and it's important the cost of doing this research is not prohibitive.

Finally, it is not clear from the consultation whether the plans include data sharing with the wider research community or the possibilities for linking administrative data with survey data. This is a major opportunity for Government and research community to develop a data infrastructure that has the potential to decrease the cost of curation by collectively improving the metadata of administrative data (i.e. map what information exists, what state it's in and whether it's been extracted for research

in the past). This documentation would also facilitate the sharing of code written for data curation and management where possible.

Question sixteen: *To ensure a consistent approach towards departments accepting or declining requests for disclosing information for research projects, should the UK Statistics Authority as the accreditation body publish details of rejected applications and the reasons for their rejection?*

We support the publishing of non-disclosive information to promote a consistent and transparent approvals process. We recommend existing metadata is made available to applicants, alongside the outcomes and other metrics of previous research applications. This should not include personally identifiable information, but would provide vital information to assist others in planning their research proposals.

Question seventeen: *What principles or criteria do you think should be used to identify research that has the potential for public benefit, or research that will not be in the public benefit?*

We would advise not being overly prescriptive about what research is in the 'public benefit'. One of the central issues in the discussion about public benefit is where securing consent is impractical. The model currently used in the Health & Social Care sector where the Confidentiality Advisory Committee (CAG) sets guidance is perhaps the best current practice in this area. Their broad approach is that the more readily it is possible to anticipate acceptance for a proposal across a broad range of 'user groups' – e.g. if one can anticipate acceptance not only from those that might benefit from the processing but also those whose data is to be processed without consent – then the clearer the indication, that the processing is, at least prima facie, in the public interest.

Alongside this we would advise of the importance of clear guidelines about ethical consent issues and the need for different tiers and more stringent requirements for more sensitive information.

Responding to the consultation

Your details

To evaluate responses properly, we need to know who is responding to the consultation and in what capacity.

We will publish our evaluation of responses. Please note that we may publish all or part of your response unless you tell us (in your answer to the confidentiality question) that you want us to treat your response as confidential. If you tell us you wish your response to be treated as confidential, we will not include your details in any published list of respondents, although we may quote from your response anonymously.


Name (optional): Maria Sigala

Position (optional): Senior Policy Manager

Organisation name: Economic and Social Research Council with input from the Medical Research Council, Engineering and Physical Sciences Research Council,

Address: Polaris House, North Star Avenue, Swindon, SN2 1UJ

Email: 

Telephone (optional): 

Would you like us to treat your response as confidential?*

If you answer yes, we will not include your details in any list of people or organisations that responded to the consultation.

Yes No

Is this a personal response or an official response on behalf of your organisation?

- Personal response
 - Official response
-

If you ticked “Official response”, please respond accordingly:

Type of responding organisation*

- Business
 - Charity
 - Local authority
 - Central government
 - Wider public sector (e.g. health bodies, schools and emergency services)
 - University or other higher education institution
 - Other representative or interest group (please answer the question below)
-

Type of representative group or interest group

- Union
- Employer or business representative group
- Subject association or learned society
- Equality organisation or group
- School, college or teacher representative group
- Other (please state below)

_____**RESEARCH FUNDERS**_____

Nation*

- England
- Wales
- Northern Ireland

Scotland

Other EU country: _____

Non-EU country: _____

How did you find out about this consultation?

Gov.uk website

Internet search

Other

Took part in INVOLVE's open policy making process _____

May we contact you for further information?

Yes No

Introductory Statement

1. The Economic and Social Research Council (ESRC) is the UK's largest funder of research on economic and social issues. We support independent, high quality research which has an impact on business, the public sector and the third sector and we have been at the forefront of developing the UK's social science data infrastructure for the last 40 years. The ESRC has led a collective response to this consultation on behalf of the other UK Research Councils, notably the Medical Research Council (MRC) which seeks to improve human health and support economic growth through supporting the delivery of world class medical research, and the Engineering and Physical Sciences Research Council (EPSRC) which funds research and training in engineering and the physical sciences, investing in a broad range of subjects - from mathematics to materials science, and from information technology to structural engineering. This response has also been informed by the views of key social science academics, funders and data infrastructure investments funded by ESRC.
2. Big Data has the power to transform public and private sector organisations, drive research and development, increase productivity and innovation, and Research Councils have made significant investments in data infrastructure to exploit the richness of data collected by government departments, businesses and other organisations. This data represent a significant and valuable resource which can be used to the mutual benefit of academic research, organisations and society as a whole. A number of Research Council investments highlighted in this response are relevant and can be valuable to enabling the improved use of administrative data under new legislative powers.
3. As highlighted by the work of the Administrative Data Taskforce (ADT) led by the ESRC in 2012 ¹, valuable scientific research with potential wider benefit to society undertaken by the UK academic research communities is often delayed, at times cut short, or never attempted in the first place, as a result of the complex legal landscape in data sharing among public authorities and the behaviours and culture of data sharing within public authorities. In response to the findings of the ADT, in October 2013, the ESRC was awarded additional capital funding from the UK government, as part of the 8 great technologies investment in Big Data and Energy Efficient Computing, and invested over £40 million to establish four Administrative Data Research Centres, one in each country, and an Administrative Data Service to these. These collectively

¹ The Administrative Data Taskforce Report (2012) <http://www.esrc.ac.uk/research/our-research/administrative-data-research-network/administrative-data-taskforce-adt/>

make up the Administrative Data Research Network (ADRN). The overarching aim of the ADRN is to enable safe access to linked and de-identified data for social and economic research of societal benefit. The ADRN Board was set up in partnership with the UK Statistics Authority, which independently governs the network and reports directly to UK Parliaments.

4. Further to investing in the ADRN, and as part of its Big Data Network, the ESRC also established Data Research Centres that seek to access data from the private sector and local government for innovative research of societal impact in social sciences and economics. These are: the Consumer Data Research Centre, the Urban Big Data Centre, and the Business and Local Government Data Research Centre. The ESRC continues to fund the UK Data Service to provide access to a wide range of data resources, and provide key support to these Data Research Centres. The ESRC together with the MRC have committed over many decades to fund longitudinal studies including the Centre for Longitudinal Studies cohorts, the 1946 birth cohort, and the household panel survey, Understanding Society. Such data, collected with consent from participants, has the potential to be further enriched through linkage with administrative data.
5. In the past three years, the MRC has invested over £100 million supporting research, infrastructure and building capacity in interrogating, analysing and linking large and complex health and social data. This includes the establishment of the Farr Institute of Health Informatics Research, a UK-wide research collaboration involving 21 academic institutions and health partners in England, Scotland and Wales. The Farr Institute is committed to delivering high-quality, cutting-edge research using 'big data' to advance the health and care of patients and the public. The MRC's data investments are underpinned by policies that both encourage data sharing and aim to safeguard the privacy and confidentiality of personal information held within these data and associated biological samples.
6. The MRC and the ESRC have a shared commitment to support health-related research more broadly, including inter-and multi-disciplinary approaches to health problems, and support research across the medical and socio-economic domains. Both medical and social science research can benefit enormously from access to both social and economic administrative data and routine health and social care records, to understand the broader context and interactions of between diverse experiences and contexts, and the essential knowledge and understanding it brings. The co-location of the Farr Institute and ADRC in Scotland, England and Wales maximises the opportunities to share resources and best practice and initiate joint research projects, which use such linked datasets to increase understanding of issues such as health inequalities and the role of social care.

7. Although the current proposals do not include health and social care data, which are the subject of different legislative vehicles, it is important for the research communities that the current legislation creates permeable boundaries between health and non-health data for sharing. The Caldicott Principles² which protect the privacy of NHS patients and regulate other parties' use of data could be a useful reference point, particularly Principle 7, that "the duty to share information can be as important as the duty to protect patient confidentiality".
8. Central to all research is the key principle that personal data must be secure and that issues around privacy, security, ethics and the legal framework are addressed. The ESRC, along with other research funders, have undertaken studies to understand more about public attitudes to sharing data. These have informed the investments made into data infrastructure and the systems to safeguard data that have been established. The RCUK Digital Economy Theme is also working closely with the Digital Catapult to support a Network in Personal Data and Trust which aims to build and nurture a community that brings together industry, the public sector, funders, research organisations, individual researchers and innovators to support the UK in becoming the global leader in trust and responsible innovation with personal data. The next growth of the internet is likely to rely on the successful generation and management of personal data (from increasingly sensor-driven Internet of Things devices and future smart intelligent autonomous vehicles etc) and high levels of trust and confidence in these data are a pre-requisite for successful new services, which have both huge economic and social potential.
9. Furthermore, the cross-cutting Digital Economy agenda led by the Engineering and Physical Sciences Research Council (EPSRC) aligns with the UK government's agenda to design, develop, digitise and deliver better Government services, which are inclusive, reaching all sections of society. Legislation should bear in mind that uptake of digital technologies and the resultant data flows from the next generation of the Internet is changing the dynamics of the way people live, work, play and behave. Future systems design must understand this rapidly changing dynamic and focus on the needs and behaviours of individuals within the digital systems.
10. Over the past eighteen months the ESRC has worked closely with the Cabinet Office and participated through the Open Policy Making process to progress work on the "access to de-identified data for research purposes" C(i) strand

² https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/251750/9731-2901141-TSO-Caldicott-Government_Response_ACCESSIBLE.PDF

and to present the views of our research communities. The ADRN model was widely discussed by the Cabinet Office and has informed proposals for a more efficient, transparent and safer way of linking and accessing de-identified data for research purposes. We are delighted that the Cabinet Office has been able to progress these discussions to this formal consultation and are supportive of opportunities to take forward proposals into legislation at the earliest opportunity.

11. New legislation that will facilitate better access to administrative de-identified data for research purposes has the potential to be of great benefit to the research community and the data infrastructures funded by the UK Research Councils and other research funders, and can place the UK at the forefront of the international scientific landscape. Research using this data has the potential to revolutionise the effective operations of government, but can only be effective if we commit to data collection systems that are inclusive and are designed to protect personal data.
12. We welcome the overall proposal for proportionate, secure and well-governed sharing of information among public authorities, and are focussed in our response on the key strand for research, “allowing use of data for research purposes” strand (Ci). In addition to responding to Questions 15 to 17 we will highlight further areas where we consider additional clarification is required if the legislation is to achieve its goals, and so our response is broader than the three questions being posed.

Questions 15, 16 and 17

Access to data which must be linked and de-identified using defined processes for research purposes

Question fifteen: Should fees be charged by public authorities for providing data for research purposes, and if so should there be a maximum fee permitted which is monitored by the UK Statistics Authority?

Yes

No

We consider that public authorities should not be charging fees to provide data for publicly funded research that advances scientific knowledge and creates the potential for other public benefits. Such a policy leads to funds being moved from one department to another and creates operational inefficiencies in the process. It can also present a barrier to undertaking research with potential public benefit that is

not being funded by a Research Council or other grant awarding body or through a government department.

However, we recognise that there are resource implications for government departments of such data sharing and so recognise that there may be a case for a cost recovery model. Should such a cost recovery model be agreed, the fees charged should be (a) strictly driven by cost recovery; (b) capped and remain modest; (c) be subject to transparency in their calculation and imposition, including accessible interdepartmental tariffs subject to external scrutiny; (d) subject to the UK Statistics Authority taking a strong role to ensure that the cost recovery model is implemented in a fair, accountable and transparent manner; and; (e) recognising that the data for which access fees are incurred may have the potential wider benefits across government and may enable efficiencies in processes currently used by government to deliver services.

It would be important to ensure that this is workable in practice, critically that researchers can be provided with realistic cost estimates for inclusion in grant applications in advance of applying for data access, and that there is clarity about 'ownership' of or access to linked data paid for through this charging.

We are aware from data infrastructures in other parts of the UK that the cost of data extraction including the creation of good metadata is variable depending on the quality and state of the data in question, among other factors. Therefore agreeing the maximum fee should be subject to review of evidence and insights and subject to scrutiny and discussion. Maximum fees should be modest to avoid (a) increasing the cost of research projects for funders, (b) hindering the commissioning of good research, and (c) risking a public backlash if perceived that public authorities are selling and profiteering from taxpayers' data.

Question sixteen: To ensure a consistent approach towards departments accepting or declining requests for disclosing information for research projects, should the UK Statistics Authority as the accreditation body publish details of rejected applications and the reasons for their rejection?

(X) Yes () No

For the purpose of transparency, it is very important that information is made available about the outcomes of requests to access and share data, whether these are approved or rejected. Our academic researchers have been concerned by the current lack of transparency and accountability when their applications to access data for research purposes from government departments are delayed or rejected.

We would suggest that there is also publication of metrics, for each public authority, on the volume of data access requests from the research and policy making communities, the volume of acceptances and rejections, the time taken to supply

data, and importantly we would seek the publication of clear criteria for access and assessment of requests. We suggest the publication of details for rejected applications, including crucial information on the reasons for such rejections, need to be carefully considered but with sensitivity to avoid disclosing personally or commercially sensitive information,.

Careful consideration should also be given to allowing the right to appeal by researchers if their request is rejected. The ADRN Approvals Panel has already such a process in place.

We consider that including such measures would provide a further incentive for public authorities to share their de-identified data for research purposes.

Question seventeen: What principles or criteria do you think should be used to identify research that has the potential for public benefit, or research that will not be in the public benefit?

The ESRC agrees with the principle that the use of administrative data for research purposes has the potential for public benefit. As discussed during the open policy making process, it is not appropriate to present a single definition of public benefit in primary legislation that would apply across all public authorities. The NHS Act 2006 makes reference to and yet avoids defining public interest, recognising that public interest is often tied up to the specifics of a case. We would support such an approach in this legislation.

It is also important to note that “potential for public benefit” should not be misread as “hypotheses driven research only”. Research that is designed to test a specific hypothesis (e.g. the link between an environmental occurrence and behaviour) may appear to be of greater public benefit than research that is collecting and linking data to improve a dataset, or to explore what data can tell us and what new hypotheses could be created. Both have the potential for public benefit.

Since definitions of public benefit may vary among public authorities, funders and other organisations, we consider that it is important that the legislation does not attempt to define public benefit as this has the potential to create further barriers to data access and use, and multiple definitions will add uncertainty and complexity to the landscape. Instead, high level principles surrounding public benefit should be presented so as to frame this safeguard. Such principles may include transparency including the imperative to make research findings publicly available; proportionate consideration of risk and benefit; lay representation on access committees; ethical review; and research that contributes to scientific knowledge and policy.

A current approach that might be of interest is used by the ADRN is their Approvals Panel which has two lay members and information about membership, meetings and its terms of reference are available on the ADRN website

<http://www.adrn.ac.uk/application-process/approvals-panel/> . Another model that

might be of interest is to consider the national importance of the research, an approach used by EPSRC in its peer review assessment of research, also used by all Research Councils. By 'national importance' EPSRC is considering how the research contributes to, or helps: maintain the health of other research disciplines; address key UK societal challenges; current or future UK economic success and / or enables future development of key emerging industry(s); national strategic needs by establishing or maintaining a unique world leading research activity (including areas of niche capability); and its fit with and complements other UK research already funded in the area or related areas, including the relationship to the EPSRC portfolio and EPSRC's strategy.

However this progresses, we would strongly recommend that the UK Statistics Authority, as the accreditation body for research, focuses on high level principles and utilising existing processes in order to ensure that the research is ethical and of no harm to individuals, rather than seek to define public benefit.

Other comments on the “Access to De-identified Data for Research Purposes” C(i) strand

(a) The limitations of a legislation giving 'Permissive' power

We note that the legislation is permissive, rather than compelling public authorities to share data. It is important in order to support data sharing for public benefit to ensure the timely access to and security of data sharing. Stronger incentives will be needed to encourage data sharing in a timely manner by government departments and further resources (including investing in upskilling staff and improving internal data infrastructures) should be put in place to further facilitate the work of the public authorities to share data with others.

Although the introduction of a cost recovery model may help to ease resource related constraints to data sharing across government, it does not address or necessarily incentivise the cultural change needed to enable better data sharing within government. Further efforts will be required, beyond the economic return from fees, to address barriers to data sharing.

(b) Ensure that boundaries to enable sharing of data are permeable, in particular with health data

The consultation document refers to the current fragmentation of the data sources, governance and arrangements, and the complexities arising. These issues apply to health data too. Furthermore, health data are treated differently in legislation, for

some important reasons, but concerns regarding sensitivity, privacy and confidentiality relating to health data may apply to other types of administrative data too. For this reason, we suggest that the concept of a spectrum or continuum may be more useful than a dichotomy where health data is treated as a separate category.

It is important to avoid creating underlying functional differences in access between two types of routine data. In particular, much research crosses the two domains, and specifically requires linkage of health and other administrative data, to each other or to other data such as survey data. It is unclear how this will be feasible under the proposed legislation and will it be facilitated. Although the current proposals do not include health and social care data, it is important for the research communities that the current legislation enables the flexibility to accommodate linkages with health data, 'permeable boundaries'.

Both medical and social science research can benefit enormously from access to both social and economic administrative data and routine health and social care records, to understand the broader context and interactions of between diverse experiences and contexts. It is important to recognise and support this inter-disciplinarity and the essential knowledge and understanding it brings.

The role of the UKSA as the accrediting body and its powers to delegate functions to other bodies will be important and it is also key such bodies have appropriate experience and are suitably qualified. Although health data is specifically excluded from the remit of the proposals, reciprocal approval bodies should have a clear and transparent process of alignment to prevent uncertainty and the potential for applications being frequently referred from one approvals' body to the other.

Furthermore, UKSA, as the proposed accrediting body, would need to be adequately resourced to fulfil its functions and avoid delaying the progress of research and data access requests.

(c) Include linkage with consented survey data

The linkage of longitudinal study survey data to routine administrative data including health records can bring enormous benefits for research. Many of the UK's longitudinal studies, including birth cohorts, in particular those supported by MRC and ESRC, span medical and social science. The longitudinal studies have particular needs in terms of linkage to routine data, including that the data and linkage need to be consistent over time or to support continuity, and typically linked data would be made available to other future research, meaning that it would be curated, stored for future use, and shared. It would be helpful to clarify whether the proposed legislation will allow and benefit this. We would support inclusion within new legislation of powers to improve the access to administrative data for linking with consented data.

(d) Clarity about private sector access

We would welcome a clear statement whether private sector organisations would have access to data. We are alert to discussions about access to data held by the Office for National Statistics that will lead to revisions to their current Approved Researcher Status, and in light of the discussions we would recommend that private sector organisations are not explicitly excluded for new legislation. Private sector organisations increasingly provide services for government, have research and analytical expertise, and are in some cases tackling similar challenges to public sector organisations. So the focus should be on enabling research with the potential of public benefit. We suggest access to private sector users could be undertaken on a case by case basis via the researcher accreditation process, and with principles such as transparency which require publications of findings resulting from utilising data in place, that would offer safeguards against access for purely commercial gain.

(e) Clarity about linked versus single data

The legislative proposals focus on data sharing in the context of linking across two or more datasets, for reasons set out in the consultation and discussed in the Open Policy Making process. However, we would like further consideration and clarity whether access to single datasets would also be improved as a result of this legislation as such research would equally benefit from such access.

(f) Access by UKSA to identified data for the purpose of producing official statistics and research

We welcome the Cabinet Office proposals for the ONS to access detailed administrative data from across government and businesses, to provide more accurate, frequent and timely statistics and to update how the census is managed, instead of relying on surveys.

We also anticipate that the UKSA gaining access to more data from commercial sources could assist ESRC funded data centres such as the Consumer Data Research Centre in their own work using this data for research purposes.

(g) Avoid over-specifying a technical solution for data linkage

The current processes for data linkage and in particular the Third Trusted Party (TTP) model was discussed during the Open Policy Making process with participants broadly accepting that it was secure and tested through existing infrastructure like the ADRN. However, the pace that technologies around Big Data are advancing will result in methodological and technological advancements and so the current models becoming obsolete or inefficient. . We would welcome changes in the draft legislation that focus on the principles of and conditions for safe data linkage rather than specifying a particular approach to data linkage.

(h) The de-identified data for research purposes strand as part of a wider Bill

Given that this consultation is presenting the case for data sharing for several reasonably different purposes, we are concerned that there may be confusion around some key differences between the needs of researchers which would be enabled by the de-identified strand and the needs of government in other strands that seek to legislate for the sharing of personally identifiable information. Although we note that the Cabinet Office has already undertaken significant work to address these differences, there is a danger that it may still cause confusion and endanger the progress and debate around the benefits of new legislation. As discussed throughout this process, we argue that strand C “Allowing the use of data for research purposes and for official statistics” has safeguards embedded within proposals and should be progressed with or without agreement on the other cases for data sharing presented in this consultation.

Response Template for Consultation on Data Sharing – Closing date 22nd April

Name : Piers Elias

Organisation : Independent Demographer

Contact : 

Summary of consultation questions

Improving public service delivery

1. Are there any objectives that you believe should be included in this power that would *not meet* these criteria?

Answer 1:...

2. Are there any public authorities that you consider would not fit under this definition?

Answer 2:...

3. Should non-public authorities (such as private companies and charities) that fulfil a public service function to a public authority be included in the scope of the delivering public services power?

Answer 3:..Yes and they should adhere to the same guidelines.

4. Are these the correct principles that should be set out in the Code of Practice for this power?

Answer 4:...

Providing assistance to citizens living in fuel poverty

5. Should the Government share information with non-public sector organisations as proposed for the sole purpose of providing assistance to citizens living in fuel poverty?

Answer 5:...Yes

6. Would the provision of energy bill rebates, alongside information about energy efficiency support, be appropriate forms of assistance to citizens living in fuel poverty?

Answer 6:...Rebates would not encourage citizens to reduce their usage nor seek energy efficient measures.

7. Are there other forms of fuel poverty assistance for citizens that should be considered for inclusion in the proposed power?

Answer 7:...Yes, enforce lower tariffs on pre-paid meters. Ensure all properties have the recommended amount of insulation.

Access to civil registration to improve public service delivery

8. Should a government department be able to access birth details electronically for the purpose of providing a public service, e.g. an application for child benefit?

Answer 8:...Yes. Can't believe you can't already do this.

9. Do you think bulk registration information, such as details of all deaths, should be shared between civil registration officials and specified public authorities to ensure records are kept up to date (e.g. to prevent correspondence being sent to families of a deceased person)?

Answer 9:...Yes.

Combating fraud against the public sector through faster and simpler access to data

10. Are there other measures which could be set out in the Code of Practice covering the proposed new power to combat fraud to strengthen the safeguards around access to data by specified public authorities?

Answer 10:...

11. It is proposed that the power to improve access to information by public authorities to combat fraud will be reviewed by the Minister after a defined period of time. This time will allow for pilots to be established and outcomes and benefits evaluated. How long should the Fraud gateway be operational for before it is reviewed?

Answer 11:...

Improving access to data to enable better management of debt owed to the public sector

12. Which organisations should Government work with to ensure fairness is paramount when making decisions about affordability for vulnerable debtors who owe multiple debts?

Answer 12:...

13. How can Government ensure the appropriate scrutiny so pilots under the power are effectively designed and deliver against the objectives of the power?

Answer 13:...

14. It is proposed that the power to improve access to information by public authorities for the purpose of better managing debt owed to government will be reviewed by the Minister after a defined period of time. This time will allow for pilots to be established and outcomes and benefits evaluated. How long should the debt power be operational for before it is reviewed?

Answer 14:...

Access to data which must be linked and de-identified using defined processes for research purposes

15. Should fees be charged by public authorities for providing data for research purposes, and if so should there be a maximum fee permitted which is monitored by the UK Statistics Authority?

Answer 15:...Fees should be permitted to cover the administrative costs of providing the data - as ONS currently do for ad hoc work. UKSA could provide guidance on fees but if a maximum is set, then it may not be viable so would depend how much work is involved -- if a maximum were set, then public authorities may refuse to undertake the work.

16. To ensure a consistent approach towards departments accepting or declining requests for disclosing information for research projects, should the UK Statistics Authority as the accreditation body publish details of rejected applications and the reasons for their rejection?

Answer 16:...UKSA publishing rejected applications wouldn't ensure a consistent approach. Rather the UKSA should set guidelines on how to assess requests. It should also publish all requests - accepted and refused. This will help other organisations see what is already available and what sort of requests are being rejected.

17. What principles or criteria do you think should be used to identify research that has the potential for public benefit, or research that will not be in the public benefit?

Answer 17:...Anything that saves time or money for the benefit of improving life chances, improved health, better education, particularly for the more disadvantaged members of society.

Access by UK Statistics Authority to identified data for the purpose of producing official statistics and research

18. Is two years a reasonable maximum period of time for the duration of a notice for the supply of data to the UK Statistics Authority for the purposes of producing National and official statistics and statistical research?

Answer 18:...No idea.

19. If your business has provided a survey return to the ONS in the past we would welcome your views on:

(a) the administration burden experienced and the costs incurred in completing the survey, and

Answer 19a:...N/A

(b) ways in which the UK Statistics Authority should seek to use the new powers to further reduce the administrative burdens on businesses who provide data to the ONS for the purposes of producing National and other official statistics.

Answer 19b:...N/A

20. What principles and factors should be considered in preparing the Code of Practice on matters to be considered before making changes to processes that collect, store, organise or retrieve data?

Answer 20:...1. Impact on disclosure. 2.Does it speed things up without additional risk or additional cost.

Full copy of Response Document and link to background papers.....

The response document – See pages 37-38 for the questions they are seeking responses on and in particular page 38 for questions 15 to 20...these are included above.

[https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/503905/29-02-16_Data_Legislation_Proposals - Con Doc - final 3 .pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/503905/29-02-16_Data_Legislation_Proposals_-_Con_Doc_-_final_3.pdf)

The background papers are available here...

<https://www.gov.uk/government/consultations/better-use-of-data-in-government>

We welcome the Cabinet Office consultation on Data Sharing and the open policy making approach that has been adopted in developing it.

We have the following general points to make about the consultation.

First, any policy in this area needs to distinguish clearly between data sharing and information sharing (and indeed knowledge sharing). In our experience these are rarely the same thing and collapsing the terms or using them interchangeably leads to confusing assumptions regarding the nature of what is being shared, for what purpose and with what intent. Data (plural) is rarely meaningful beyond very specific boundaries and relationships and may be analysed or processed in a number of ways. Information implies the act of informing someone of something, and implies a much more defined interpretative frame if the information is to 'make sense'. We have tried to summarise this point in the title of our ESRC funded Seminar Series in the phrase "what else needs to be shared (or not shared) when we share information?" Because these fundamental distinctions are not clearly made in the consultation document, we are concerned the remaining edifice of argument rest on shaky foundations.

Second, the implicit model in the consultation is an individual citizen in a range of transactions with a monolithic state and/or its agents. By its very nature this is a reductive model focusing on just one form of relationship between government and citizen - discrete transactions. Drawing on our empirical work exploring how the various parts of public services use information (or not), we would supplement and extent this view by emphasising the relational nature of information in the interactions between the citizen and the state and the ways in which informational transactions are often experienced by citizens as interdependent, tangled and linked into (often confusing) narratives (see e.g., Dawes, et al. 2009).

Third, in our empirical research we have identified a much richer set of interactions between the Government, its agents and those who live and work in the UK than is envisaged in much of the consultation document. For instance our work on families (Cornford et. al 2013a) shows seeing the family as a unified (or unifiable) and stable category or specific data item belies the complex nature of family lives. Collecting data and information about families tends to be service specific and the nature of familial relations are represented accordingly, leading to inconsistencies in the nature of what is shared about families. We could make similar points about the nature of information sharing in the lives of older adults, especially those with reduced capacity for decision making (whether chronic or acute), or military veterans (who may or not want to be identified as such) (Wilson et al. 2015). These are more complex situations, in which there is an information sharing problem on the citizen sides of the transaction as much as on the public services side. The ability for local public services to be delivered for those most in needs often relies on the ability to know of, and to know about, such individuals in this wider context (what we have called the 'view from somewhere' – Wilson, et al. 2011, Cornford, et al. 2013b). In these contexts, information is likely to distributed around the various agencies (formal and informal - public/private and voluntary) and across family and other social networks (McLoughlin, Wilson and Martin, 2013).

Fourth, with regards to the specific questions regarding contexts including Fuel Poverty, Civil Registration, Combating Fraud, Debt and Research it is our view that Government agencies and those delivering public services should have the ability to make referrals to each other subject to the appropriate consents, communication and recourse mechanisms being in place. The emphasis on particular situations for data sharing implies an extension of the 'Gateway' model described in the Law Commission review as a significant issue ripe for policy simplification: "The large number of legislative gateways, spread across primary and secondary legislation, is difficult to navigate and creates complexity. There are express and implied gateways, permissive and mandatory gateways, gateways which restrict use or onward disclosure and gateways which do not." (Law Commission, 2014: 73). The consultation seems to have interpreted this statement to mean that the number of gateways should be reduced. We would argue that a more radical interpretation is required, one that questions the model of gateways as the primary way of providing controlled, accountable information sharing in public services and that addresses the mechanisms that have generated a plethora of gateways in the first place (see McLoughlin, Wilson and Martin, 2013). If these more fundamental architectural issues are not addressed, we suspect that the success of attempts to reduce the number of gateways and simplify their procedures will be short lived.

Fifth, our recent review "Information sharing is easy to say, but harder to do well" (Wilson and Gray, 2015) for the Centre of Excellence for Information Sharing made the following additional points about current state of knowledge (academic and practitioner) around information sharing in particular but which have relevance to the data sharing issue.

- "That the information sharing literature is dispersed over a wide variety of academic disciplines, policies, and professional contexts. The term is used in different ways, and is taken to mean different things, in all these contexts. Previous experience in health, care and education services, where dispersed public services are managed using approaches such as information governance, suggest that it is highly unlikely that a 'one size fits all' or purely bureaucratic or procedural approach to information sharing, based on the requirements of one organisational perspective, is going to 'solve the problem' on its own."
- "Dealing with complexity requires the development of appropriate tools and interpretative skills (both individually and collectively) for those involved in public services. For example, individuals and organisations need to understand the provenance of information that is being shared – who published what, about whom, when, with what context, and with what authority – and they need the skills and tools to make judgements based on all of these issues."

Improved – simplified or clarified – data governance of the kind envisioned in the consultation document, while important, is therefore unlikely to result in improved sharing, and therefore more economical, efficient and effective public services, as long as other issues are not addressed at the same time. To make improvements the need for information sharing need to be built into public service professionals' role structure (including training and job descriptions), it needs to be reflected in reward and incentive structures, and it must be reflected in the patterns of relationships and routines – the professional culture of

public services (Richardson and Asthana, 2006). Each of these elements needs to be adequately resourced. Without addressing these issues, even well designed changes to data and information governance are likely to fail to make a difference, leading to counterproductive cynicism.

Finally we would observe that previous government policy on co-production in public services has envisaged a more active citizen being more responsible and engaged in the production of services. This potential can only be realised if the ways in data and information is conceived in public and social policy is represented as an integral part of that vision. It is possible to describe a more sophisticated approach which could lead to better service coordination, practitioner confidence, information sharing behaviour and service delivery: local service communities could work together to improve their information sharing relationships (internally and externally), in parallel with improvements in systems and resources. In the context of the government's devolution agenda for England, the issue of such local 'interpretative communities' (Cornford et al., 2013) and the local governance of information sharing (Wilson et al., 2013) are increasingly important. Central policy is only ever likely to be a part of such an approach.

*Rob Wilson (Professor, Newcastle University), James Cornford (Senior Lecturer, University of East Anglia), Sue Richardson (Lecturer, Bradford University), Sue Baines (Professor, Manchester Metropolitan University)**

- Cornford, J., Baines, S., and Wilson, R. (2013) Representing the Family: how does the State "Think Family"? *Policy and Politics* 2013, **41**(1), 1-18.
- Cornford, J., Wilson, R., Baines, S. and Richardson, R. (2013) Local Governance in the new information ecology: The challenge of building interpretative communities, *Public Money and Management* 2013, **33**(3), 201-208.
- Dawes, S.S., Cresswell, A.M. and Pardo, T.A. (2009) From "need to know" to "need to share": Tangled problems, information boundaries, and the building of public sector knowledge networks, *Public Administration Review*, **69**(3): 392-402.
- Law Commission (2014) Data Sharing between Public Bodies A Scoping Report. London [http://www.lawcom.gov.uk/wp-content/uploads/2015/03/lc351_data-sharing.pdf]
- McLoughlin, I., Wilson, R., and Martin, M. (2013) *Digital Government at Work: A Social Informatics Perspective*. Oxford: Oxford University Press.
- Richardson, S. and Asthana, S. (2006) Inter-agency Information Sharing in Health and Social Care Services: The Role of Professional Culture, *British Journal of Social Work* **36**(4): 657-69.
- Wilson, R. and Gray, A. (2015) *Information Sharing: Easy to say harder to do well*. Leicestershire, UK: Centre of Excellence for Information Sharing, 2015.
- Wilson, R., Baines, S., Hardill, I., and Ferguson, M. (2103) Information for Local Governance. Data is the solution... What was the question again? *Public Money and Management* 2013, **33**(3), 163-166.
- Wilson, R., Cornford, J., Baines, S, and Mawson J. (2011) Information for Localism? Policy. Sense-making for Local Governance. *Public Money & Management* 2011, **31**(4), 295-300.

Wilson, R., Harrison, D., Gray, A., Martin, M., Jenkins, K.N. (2015) *Evaluation of Information Sharing Practice in the North East mixed economy of service provision for ex-services personnel*. London, UK: Forces in Mind Trust (FiMT), 2015.

*We are a group of academics who have been working on aspects of the data and information sharing challenge for over 15 years including work sponsored by DCLG, DH, Cabinet Office as well as Innovate UK, ESRC and EPSRC. We work closely with the Centre of Excellence for Information Sharing (CEIS) but we would like to point out that our views are not the views of the Centre. This submission has been made on behalf of the ESRC funded Seminar Series on Information Sharing but again are not the views of the ESRC. For more information about the series see www.ncl.ac.uk/kite/esrc_seminars/