Summary of civil society and public sector policy discussions on data use in government

Contents

This document is a summary of civil society and public sector policy discussions on data uses in government. This document captures the discussions that have taken place. In circulating this document we are not seeking to open up new issues or re-open issues that have already been discussed.

Introduction

1. Since March 2014 civil society organisations, privacy groups, government departments, academics and the wider public sector have been collectively discussing how government can be made to be more efficient and effective through its use of data. The core focus has been to enhance the availability of high quality research and statistics from administrative data; prevent fraud, reduce error and help citizens manage debt they have with government; and ensure the right services are offered to the right person at the right time. This document sets out the current thoughts emerging from each of those areas.

2. The starting point was work already undertaken within government focused on data sharing as the solution. However, it is well known that significant sensitivities exist about where the balance is between the proportionate use of data to deliver services and maintaining people's privacy. Similarly it is an area where citizens' trust levels of government can be variable.[1] To ensure there was a shared understanding of both what government wants to achieve and the concerns of those outside of government, an open policy making approach was suggested.

3. Involve, a not for profit organisation established to improve public engagement with government, were asked by the Cabinet Office to work collaboratively on an open policy making process on this issue. Involve agreed to help facilitate the process and external engagement. Since then over two hundred organisations have been invited to participate based on their particular interest in the issues considered. The process has been open to any interested organisations to join. The number of non-government individuals and organisations willing and able to participate in the process was less than anticipated. Resourcing issues, particularly where the organisation is relatively small, has been a key factor. Groups engaged in the process include those with a specific interest in individual privacy and rights, academics, statisticians, researchers and their funders, charities, government officials and some private sector organisations. Representatives from these organisations and individuals have participated in discussions and the development of proposals in each of the themes. The scope of this new joint work was expanded to include looking at alternative solutions, not just data sharing. The task set was to examine the evidence and use the analysis to inform the design of policy proposals from it. The process was designed to ensure that all voices were heard from the outset, increasing the likelihood of balanced, successful policy recommendations being delivered.

4. Transparency has underpinned the whole process, with all work as open as possible. Key information and updates have been posted on

---

[1] Sciencewise (2014) Big data: Public views on the collection, use and sharing of personal data by government and companies.

www.datasharing.org.uk, a non-government website, to act as a repository and audit trail of the work.

5. The following key principles have underpinned the discussions:

    a. Proposals would **not** consider the building of new large and permanent databases, or collecting more data on citizens;
    b. Proposals would **avoid** the indiscriminate sharing of data within Government; and
    c. Proposals would **not** weaken the Data Protection Act.

6. The interim findings set out in this paper offer a balanced consideration of options including data sharing as a solution to the specific challenges we looked at. It is the product of a truly open collaboration between a range of public sector officials and civil society organisations and privacy campaigners. In some instances, the consensus among the group was that legislative changes along with appropriate safeguards would be the best course of action to achieve the objectives. Whilst with other issues, the consensus was that more work needed to be carried out to assess the value of interventions involving data sharing. These balanced recommendations demonstrate the value of the open policy making approach and collaborating with a range of organisations with different perspectives.

7. There remain some areas, particularly around safeguards that will need to be in place where data is to be shared, and implementation, where more work on the detail needs to be carried out. On these issues further dialogue is required, so that an appropriate way forward can be agreed before final recommendations can be made.

8. Open policy making still represents a new way of working. Taking this approach to an issue as challenging as data sharing builds on the excellent work delivered by using this approach to develop the UK's 2013-15 Open Government Partnership national action plan. This process provides further evidence of the value of working in partnership to look at some of our biggest challenges. These interim findings are a good basis for further consideration and continuing the process with a stronger focus on the detail of implementation and review.

Summary

9. The open policy making process has been extensive, with engagement taking place at individual and group levels, either for individual specific policy challenges or across the policy areas. 17 sessions were run with a large number of representatives between April and July. Representatives from both within and outside government have listened and changed their position where the case has been compelling. To ensure transparency, all key discussion documents were posted on www.datasharing.org.uk, a non-government website, and invited comments from the general public.

10. The following three specific policy challenges were explored as part of the open policy making process:

   - Research and statistics – improving the quality of statistics and enabling the availability of better evidence to inform the formulation of policy and delivery decisions;
   - Fraud, error and debt – saving taxpayer's money wasted on fraud and error and provide those citizens with multiple debts to government greater support to help manage their debts; and
   - Tailored public services – improving the tailoring of public services so that the right services are offered and provided to the right person at the right time.

11. A brief summary of the key findings of the process to date are:

Research and statistics

De-identified data

   - Representatives from both within and outside government recognised the need for public bodies to be able to link data for research purposes. Representatives were supportive of a proposal, provided data linking was carried out in a de-identified, and therefore more privacy enhancing way, using a Trusted Third Party sharing system. Trusted Third Parties, researchers, and the subject of the research would all have to be accredited under a system established through legislation by an oversight body. Research would have to demonstrate that it is for the public good. Extensive consultation via the open policy making process led to consensus on the aims and proposed powers.

   - It was agreed that further work would need to be carried out to refine the requirements for accreditation, the oversight body as well as the implementation process, and any exclusions from a future power, such as health services bodies.

<u>Identified data</u>

- Participants in the open policy making process considered a specific proposal from Government aimed at improving access by public authorities to identified data for research and statistical purposes. The proposal is to enable public authorities to disclose data to the Office for National Statistics (ONS) in order for them to carry out the executive functions of the UK Statistics Authority (Authority) to provide statistics that serve the public good. This would reduce the burden on businesses and other respondents by reducing the cost of surveys as well as time taken to by respondents to complete them; improve policy making decisions based on research and statistics by strengthening the evidence base for policies; improve the quality of statistics, while preserving the privacy of data subjects and ensuring that data are used appropriately, by ensuring that safeguards are embedded in the process. It was agreed that changes to legislation would be required to meet the identified objectives.

- The group agreed that alternative options for the scrutiny of proposals for disclosure of administrative identified data to ONS would require further work and continued engagement through the OPM process.

<u>HMRC general and aggregated and de-identified data</u>

- Participants in the open policy making process considered a Government proposal to reduce the restrictions around the disclosure of less sensitive general, aggregated and de-identified HMRC data for public benefit. Legislation limits the circumstances in which HMRC may share information. Other public bodies are not so restricted, and thus this proposal helps achieve greater equality for HMRC to contribute to the consideration of wider government initiatives and to academic research than it currently can. Representatives from within and outside government agreed with the aims and the rationale for legislative change.

- The OPM process however concluded that further work was required, specifically on developing alternative options on how HMRC could use its data more effectively by, for example, delivering benefits beyond HMRC's own functions, in balance with protecting confidentiality.

<u>Fraud, error and debt</u>

- Representatives from both within and outside government called for more robust evidence to be gathered on a range of fraud, error and debt issues (from the scale of the problem to the value of different types of intervention) from which assessments of potential options could be made. To address this evidence gap, steps proposed to be taken include:

- o a set of surveys to gain better insight into the public understanding of what the Government currently does with data to tackle fraud, error and debt and the tolerance that the public has for further sharing;

- o the development of a set of case studies where data sharing to reduce instances of fraud, error and debt has and has not been successful in achieving its objective, with a clear understanding of what has incentivised good sharing and what has driven poor sharing;

- o a set of trials and pilots of new ways of improving data use in order to understand the comparative value of different types of intervention. Two pilots have been suggested so far: one that tests an approach that seeks to filter out low-risk citizens so that more resource can be focussed on those with a higher risk of committing fraud; and a second that asks citizens to validate information held about them when decisions are made in order to correct any errors; and

- o case studies and pilots will also provide greater insight into the overall costs and benefits of intervention.

- Consensus between representatives from Government and civil society groups was that measures to address problems related to error were best addressed through better communications between agencies as well as between agencies and citizens rather than specific data sharing measures. It was agreed that validating with citizens the information held by organisations when making a decision was the preferred option. Any mass validation approaches would be disproportionate use of data and not viable as a suitable proposal.

Tailored public services

- Participants in the OPM process agreed in the value of exploring data sharing to support the delivery of more tailored public services. To test the opportunities and barriers a number of objectives were considered, including:

  - o reducing the number of 16-19 year olds not in employment, education or training;
  - o improving health, education and employment outcomes for families with multiple disadvantages; and
  - o supporting the improvement of health outcomes amongst pensioners.

- The initial proposal was for specific gateways to address specific issues. Through an iterative policy development process, the group have developed and reached agreement on a broader but constrained power, which active participants from civil society organisations then pushed to broaden the power slightly further from target groups to supporting policy objectives. The current thinking is to introduce a permissive power for defined public agencies to share data with defined public agencies for the purposes of improving the delivery or targeting of public services where it supports the

achievement of the defined policy objective and the direct benefit of the citizen.

- Further work is required to agree how best to describe the outcome in a way that constrains but also future proofs.

- The group has agreed that further work is required to explore potential issues around derived data as a result of the proposed permissive power and whether additional safeguards need to be put in place. It was also agreed that an assessment of potential burdens would need to be carried out as well as further work to identify case studies against which the powers would be tested.

Safeguards

12. Throughout this process we have considered and challenged whether data sharing and legislative solutions are required to achieve the desired outcomes or whether other approaches could be taken. We believe these initial recommendations reflect this approach. They adhere to our principles that we would not develop solutions which require the building of large permanent databases or collecting more data on citizens, or weaken the consumer or privacy rights provided by the Data Protection Act and other regulations.

13. Where a solution, which requires changes to data sharing legislation, has been recommended, it is supported by measures that provide appropriate safeguards to protect the privacy of citizens. The need to be transparent has been at the core of the discussions and as a result options include making privacy impact assessments available for public scrutiny. In developing these proposals we have sought to balance a consistent approach across the different areas with the necessary tailoring to ensure that the unique features of each area are addressed appropriately, informed by the broader framework within which they fit. This has led to a few key differences in approach being taken across the three strands.

Devolution

14. Cabinet Office is currently engaging with officials from devolved administrations in Wales, Scotland and Northern Ireland as they begin their own open policy making processes on the three strands to determine whether there is interest in applying the policies at all, identically, or with particular differences. Devolution issues are complex because, for example, whilst policy on data sharing may be delegated, key datasets such as the DWP Customer Information System (CIS) covers England, Wales and Scotland, but not Northern Ireland.

15. We will consider findings as they emerge and identify where there is scope for cooperation to achieve agreed outcomes. The final summary document will capture and summarise options identified by the devolved administrations.

Part 1 – Why are we doing this?

16. The public sector faces a number of difficult challenges to deliver better services more efficiently. Private and public sector technology enabled services such as online banking, shopping and registering to vote have improved the speed and way services are delivered and as a result customer's expectations have increased. Furthermore, public bodies continue to work within tight spending constraints and must continue to reduce waste and find the most efficient way of delivering their objectives in a way relevant to the world around us.

17. Accurate and timely data underpins the delivery of many modern services whether public or private. Good data are critical to help inform the decisions we make throughout the lifecycle of public service delivery. Population statistics and other data provide the evidence, which skilled analysts and policy officials use to inform the policy formulation process and appraisal of options. Data are also used to inform the key operational decisions that ensure the right services are offered and delivered to the right citizens at the right time. They can also be used to reduce waste such as instances where multi-agency cooperation can help identify when taxpayer's money is lost through fraud as well as reduce duplication of investigative and administrative functions across agencies.

18. A good example of what can be achieved by way of effective data sharing to deliver an outcome with a public benefit is the recent work by the Cabinet Office and the Department for Work and Pensions (DWP) on Individual Electoral Registration (IER). The introduction of IER brought in measures to reduce electoral fraud, which utilised data matching to confirm the legitimacy of applications to register to vote. The process is designed so that an elector's personal identifiers submitted in an application are transmitted securely via an application programming interface (API) for data matching against existing records to confirm their identity. The process does not aggregate or bring citizen's data together in a new way, but safeguards the privacy of the data by only providing an indicator to the local authority whether the elector's details have been matched or not, which will then prompt further action as necessary. IER and the adopted approach to handling data was consulted upon widely, extensively debated in Parliament, and is supported by all main political parties and non-party bodies concerned with the running of elections, as well as the Information Commissioner.

19. Data sharing between public bodies provides a number of challenges. The first of these is the complex legal landscape. The Law Commission scoping report, *Data Sharing between Public Bodies*, describes how the law surrounding data sharing is complex, with powers to share data scattered across a very large number of statutes which may be set out expressly or implied. The report identified that

there are problems in practice and that there are differing interpretations of the law governing the sharing of data.

20. Understanding the complex legislative landscape around data sharing can be difficult. There have been instances where public bodies have decided to introduce specific statutory powers where data sharing is required rather than understand what existing legislation permits. The process of bringing in such explicit powers can cause significant delays and additional cost due to pressures resulting from parliamentary process. Such delays to the sharing of data can prevent early intervention or action for those most at risk.

21. The second key challenge involves citizen concerns about data sharing. Citizen trust can be built through accountability via greater transparency of the data sharing process and control via ensuring that the protection of privacy rights is fundamental to any proposed data sharing regime. Consultation carried out by the Law Commission during their work on the scoping report on data sharing between public bodies indicated that there are a wide variety of public attitudes to data sharing and varying levels of public trust. Striking the appropriate balance between privacy and the public interest is a key consideration in establishing any solution involving the sharing of data.

22. After some initial background work on data sharing issues, an open policy making process was launched in March 2014. Three specific problem areas were identified to be explored through the open policy-making process, with the objective of developing policy and legislative solutions where appropriate. The three areas were:

- Research and Statistics - making it easier for accredited researchers to access linked administrative data sets in accredited secure data access facilities, and speeding up the access the Office for National Statistics (ONS) has to data for statistical information;
- Fraud, Error and Debt - looking to build on and improve the way we use identifiable data across boundaries to prevent and reduce instances of fraud and error, and help citizens to better manage debt with government in a more holistic way; and
- Tailored Public Services – maximising the benefit of data already held by public bodies to deliver public services tailored to individual needs.
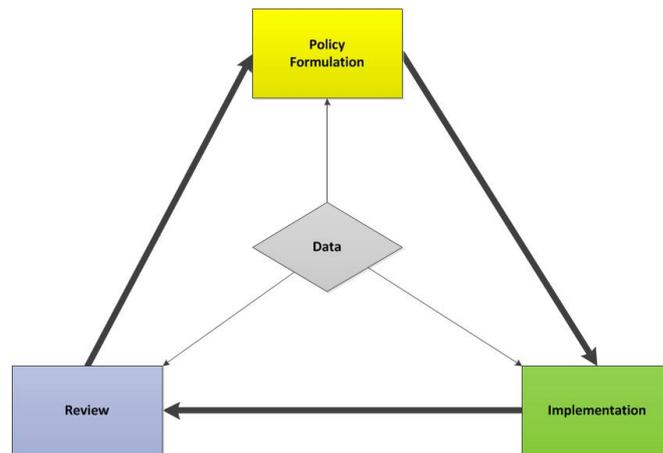
23. The three areas identified represent different stages of the policy cycle and opportunities where better data could improve the services offered for public benefit. Research and statistics provides the evidence base, which informs policy formulation and operational decisions. Tailored public services concerns how policy can be better implemented so that the front-line has the information required to offer and deliver the right services to those in need. Fraud, error and debt is a good indicator of the challenges of implementing and reviewing the success of policy to reduce waste and help citizens to better manage debt with government in a more holistic way.

Part 2 – Joint Policy Options

1 – Research and Statistics

A) Linking of de-identified data for research for the public benefit

The situation

24. The aim of the proposal is to ensure that public bodies (apart from health services bodies) across the UK (if this is agreed by the Devolved Administrations) are able, if they so wish, to engage, for the purposes of research, in the process of linking two or more datasets from two or more data controllers in a de-identified and therefore privacy enhancing way using a Trusted Third Party (TTP) Sharing system which has been accredited under the legislation. Accreditation would be by linkage using a Trusted Third Party system ensures that identifying data and payload data are always kept separate, so that researchers never see identities and the Indexer sees nothing but identities. A glossary is attached at Annex A. A diagram of how the TTP sharing could work is set out at Annex B.

25. The December 2012 report from the Administrative Data Taskforce, *"The UK Administrative Data Research Network: Improving Access for Research and Policy"* recommended that "primary legislation should be sought to provide a generic legal gateway for research and statistical purposes that enables efficient access to, and linkage between, administrative data held by different government departments, agencies and other statutory bodies."

Evidence for and against change

26. Examples from government, academia and the third sector of the successes of accessing linked administrative data sets are contained in Annex C, together with the difficulties experienced in gaining access to linked data sets. This suggests that easier access to data would assist research, within government and outside, including for example: improving the family justice system; examining the drivers of productivity growth; energy saving and consumption; offending and employment; the creation of indices of deprivation; design of a local council tax scheme for vulnerable groups; housing and planning policy; and enabling the National Council for Voluntary Organisations to inform policy around fundraising. The policy is aimed at research that is for the public good, of which the foregoing are merely examples.

27. The ESRC and ONS commissioned research from Ipsos Mori published in 2014. The findings suggest that the public would be broadly happy with administrative

data linking for research projects provided (i) those projects have social value, broadly defined (ii) data is de-identified, (iii) data is kept secure, and (iv) businesses are not able to access the data for profit.

28. Provided the safeguards are properly implemented by the overarching accrediting authority, no evidence against this change has been presented to the Open Policy Making Group. Concerns that de-identification "does not work" are sometimes expressed in the media, but this is not borne out in the Report of the Administrative Data Taskforce and other recent reports. The possibility of de-identification through the use of privacy enhancing techniques is also acknowledged by the Information Commissioner's Office.

Options identified and appraised

29. The following options were identified and appraised (full details are set out at Annex D):

   a. a broad generic power: a general power for any and all public authorities to disclose identified data to each other for research and statistics; at the Open Policy Making Plenary meeting on 22 October 2014, members of Civil Society, present in person and by proxy, made it clear that this option would be a step too far and would be a red line for them. It will therefore not be taken forward.
   b. a power for a single data controller to use a safe haven to process or disclose their own data for research or statistical analysis;
   c. Do nothing; and
   d. Power for public authorities (except health services bodies) to link de-identified data for research using accredited bodies and TTP sharing (recommended approach - see below).

Recommended approach

30. A power to ensure that all public bodies (except (a) health services bodies and (b) adult social care bodies in respect of all personal data relating to service users), across the UK (if this is agreed by the Devolved Administrations) are able, if they so wish, to engage, for the purposes of research, in the process of linking two or more datasets from two or more data controllers in a de-identified (and therefore more secure) way using only accredited bodies and a particular method of sharing called a Trusted Third Party Sharing system.

31. Data linked under the powers could be used for statistics produced by a variety of Departments, as illustrated in the Annex C. This power will not cover the following:

a. health service bodies and adult social care bodies are excluded because clinicians, patients and members of the public have all expressed serious misgivings about sharing confidential health information for secondary purposes. The Department of Health is currently consulting on proposals to make new regulations under s251 of the NHS Act 2006 to make provision for Accredited Safe Havens – secure environments where this information can be processed for secondary purposes;

b. any processing of personal data that is not for research purposes e.g. processing for operational purposes;

c. processing by a single data controller; and

d. processing of identified payload data.

32. The proposals outlined in this section of the policy paper would ideally apply throughout the United Kingdom, but since data sharing is a devolved matter in relation to certain functions under the different settlements, Cabinet Office is currently engaging with officials from devolved administrations in Wales, Scotland and Northern Ireland to determine there is interest in applying the policies at all, identically, or with particular differences.

Scope

33. The Cabinet Office examined the existing position and surveyed a variety of public bodies. This revealed that many could already engage in Trusted Third Party Shares in some (but not all) circumstances. It was therefore considered unnecessary to attempt to legislate entirely for every single step in the process of Trusted Third Party Sharing.

34. Consequently the proposed legislation is intended to be restricted to covering two particular areas:

a. The first is to create a specific provision in legislation that provides any public body with the necessary powers to engage in trusted third party data shares linking de-identified data with one or more other sources for the purposes of research of any kind. However in order to recognise that many public bodies already have such vires in whole or in part, this provision will need to be carefully considered so as to only provide the powers where it is currently lacking, thus ensuring no existing vires or power is implicitly repealed.

b. Secondly given that the previous area described above is wide in both the scope of the bodies it would apply to, and in the scope of the material covered, it is appropriate that the proposed legislation would

also need to include explicit safeguard provisions to counterbalance them. Therefore it has been proposed that, in addition to the vires provision, that other provisions are made appointing a body to have oversight over a process of accreditation. The oversight body would accredit Trusted Third Parties, researchers, research and secure access facilities. The vires provision would be restricted so that it only provides vires where all the relevant bodies involved are properly accredited under the legislation. While minimum accreditation requirements could be set out in the legislation, the body providing oversight of the accreditation process would themselves develop and publish detailed standards and requirements to attain and maintain this accreditation.

35. The proposed legislation would not make any other further provision as regards Trusted Third Party Sharing, leaving the rest of the administrative process to be agreed between the parties on a case by case basis and be governed by the overarching and untouched relevant pieces of regulatory law such as the DPA. Additionally any requirements of the legislation (such as accreditation) would only apply to Trusted Third Party Sharing that was relying upon the vires provision in the legislation (by at least one of the data sources). Anybody who already had sufficient vires to undertake such sharing would be able to continue to do so if none of the sources in the share needed to use the vires, but they would not be bound by any of the requirements of the proposed legislation (such as necessity of using only accredited bodies, though they would of course remain bound by any other applicable laws (e.g. DPA/law of confidence)). Furthermore the proposed legislation will not prescribe DPA roles to the parties involved, such as data controller and data processor, although it seems likely that Trusted Third Party Indexers will generally be data processors. The parties concerned will establish this on a case-by-case basis in light of their own personal positions and the requirements of regulatory and other law generally. The expectation is that data sharing agreements will be time limited and that, in accordance with the Fifth Principle of the Data Protection Act, personal data processed for a purpose will not be kept for longer than is necessary for that purpose.

36. Based on the initial policy development in this area, previous reports on this subject and development of the proposals with internal and external stakeholders, initial definitions for key terms as regards the particular circumstances of the proposed legislative provisions have been developed. These are contained in Annex A (work on this glossary will continue). These are subject to change and will be further refined both through the open policy development process and through working with Parliamentary Counsel. The purpose of this glossary is to illustrate effectively the circumstances intended to be covered by the proposed legislative provisions but the legislation that results

may define terms in a different way, e.g. so as to ensure consistency and clarity in the law more widely.

Oversight body

37. The accreditation requirements under the legislation would be overseen by an oversight body. The identity of the oversight body may be specified in the legislation or alternatively the legislation may provide a power or procedure to appoint this role. The characteristics of an oversight body would be: independence, expertise in statistical research and analysis, and reporting directly to Parliament. The UK Statistics Authority is an example of the type of organisation that could be the oversight body.

<u>Possible safeguards</u>

38. Legislation or guidance could set minimum requirements for accreditation. Precise requirements will be determined in due course as this policy develops and is refined to a finer level of detail. Minimum requirements should be:

a) Accredited Secure Data Access Facility (ASDAF): cannot be one of the data sources in a particular data share (though a data source can be an ASDAF for another person's data in a different data share), must be fit and proper, and have sufficient procedures to ensure that researchers are only ever given access to de-identified data, and that only aggregate (i.e. not individual level records) data can leave the control of the ASDAF or be published or disclosed by researchers.

b) Accredited Trusted Third Party (TTP): could not be one of the data sources, must be fit and proper and have sufficient procedures in place to prevent data from being removed or disclosed from the Indexer contrary to the terms of the share, the requirements of the accreditation, or any legal or contractual prohibition.

c) Accredited TTP Researcher: must be a fit and proper person, conducting approved research for public benefit, having undertaken and kept up to date training on correct data handling. It is understood that the Administrative Data Research Centres are not currently planning to consider private sector research requests for data. As a matter of policy we do not intend to exclude the possibility of private bodies or persons becoming accredited researchers.

d) Accredited TTP research: research that serves the public good, which could in particular include (i) increasing knowledge about social and economic matters, and (ii) assisting in the development and evaluation

of public policy. The outcome of the research would have to be published. Note that the ADRN is establishing an Approvals Panel (Including lay membership) which will approve research if it fulfils all of the following conditions:

> a) necessary (i.e. the information does not exist elsewhere);
> b) feasible;
> c) of scientific merit (i.e. be worth asking);
> d) has assessed and mitigated privacy issues;
> e) has gone through a formal ethics review;
> f) benefits the public; and
> g) will be published.

39. The oversight body would ensure that the identity of the approved researchers and the purpose for which the research has been approved is made available to the public, together with a plain English summary of the outcome of the research.

40. We recognise that the issue of the interaction of different safe havens will have to be carefully managed. Safe havens, such as the HMRC Datalab, the Accredited Safe Havens for NHS England and the ASDAFs referred to above are, or would be, established under different powers for different purposes.

Next Steps

- Further discussion around composition and functions of the oversight body under the de-identified data proposal.
- Discussion with officials from the devolved administrations as to the extent that they would wish the proposals to apply to their territory, where powers are devolved.
- Developing detailed plans for implementation.

B) Identified data for research and statistics

The situation

41. Nothing stays the same. Society and the economy are constantly changing, bringing fresh challenges to Government. The policy responses to these challenges must be based on evidence that helps policy makers to understand underlying causes – to ensure that interventions are appropriate, properly targeted and make the best use of public funds. Evidence is also needed to monitor the effectiveness of Government policies, and to hold policy makers to account.

42. Independent, high quality statistics have a vital role to play in a democracy and, to provide these, it is important that statistical producers have access to as wide a range of data as possible. This includes access to data collected by public authorities as part of their routine business. These administrative data, particularly when they are linked and matched with data from several sources, can provide a rich and flexible source of evidence about how society and the economy are changing.

43. Currently, when a new policy issue arises, it is not easy for statistical producers in Government to gain access to information from administrative data sets that are held by other departments. Data sharing raises complex legal and policy issues which are open to different interpretations. This leads to an understandably cautious approach on the part of data owners. It can take months or years to reach agreement about whether a data source is relevant, and whether it can be shared. In the meantime, policy makers sometimes find themselves forced to make decisions without a comprehensive evidence base.

The legal framework: the Authority and ONS

44. The Statistics and Registration Service Act 2007 (SRSA) established the Statistics Board and set out its powers. The Board is now known as the UK Statistics Authority ("the Authority"). The Authority reports directly to Parliament and the devolved legislatures, rather than through Ministers. The Authority has oversight of ONS. ONS is the Authority's executive office; it is the UK's National Statistical Institute and the UK's largest producer of official statistics. For the rest of this document, ONS is referred to when describing the executive arm of the Authority.

45. The remit of the Authority, and therefore ONS, is limited by the SRSA (ss8-28) primarily to:

a. producing official statistics;

b. promoting and assisting in statistical research; and

c. providing statistical services.

46. The SRSA defines official statistics (s6(1), a definition which is broader than those statistics produced by ONS. The Act also requires the Authority to produce a Code of Practice for Statistics. The Board must keep confidential personal information held by it (s39) and its unlawful disclosure is a criminal offence.

47. Information disclosed to the Authority under the powers of access to information discussed in this document would in practice be provided to ONS, and the remainder of this document.

The Current Legal Framework for Sharing Identified Data with ONS

48. At present some departments are able to use their prerogative or common law powers to disclose information to ONS. For example the Home Office shares Border Agency data with ONS for the purpose of migration statistics.

49. The SRSA contains other powers to allow certain identified data to be provided to ONS.

50. Sections 42,43 and 44 allow certain limited types of identified data to be supplied to the Authority in respect of births and deaths and NHS registration in England and Wales.

51. Under s45 HMRC may disclose information to the Authority, but under s45(5) this may not include personal information other than for import or export statistics. This is one aspect of the SRSA that the Cabinet Office, with ONS and HMRC support, is seeking to amend.

52. Where no other data sharing gateway or power exists, or where the disclosure of information is expressly prohibited, s47 allows the Minister for the Cabinet Office to make regulations, called Information Sharing Orders (ISOs), to authorise a public authority to disclose information to the Authority to enable the Authority to carry out one or more of its functions under the SRSA (but not to provide statistical services).

53. The Cabinet Office, with the support of ONS, is seeking to amend these provisions because the s47 power is subject to significant limitations:

a. ISOs may only remove a barrier contained in a rule of law or an Act passed before the SRSA, not after 26 July 2007. As a result the ONS is unable to use an ISO to access information where the prohibition on disclosure came into force after 2007. ONS has found that, in practice, teams working on Bills since then have been reluctant to add a data sharing clause with the same effect as s.47 to their Bills, even when the departments support the principle. This is because of the potential to disrupt the passage of the Bill over what is considered to be a secondary issue. This was an issue with the Electoral Registration and Administration Bill.

b. Although ISOs can be used to create gateways where none already exist, ONS has found that the time taken to obtain agreement from the relevant departments can be considerable. Much of this is due to the need to resolve the uncertainties about whether a new gateway is necessary. Only then can work begin on establishing whether there is a sound justification for the data share. Experience has shown that this can be a lengthy process taking many months.

c. Reflecting the general caution around data sharing, and specific concerns about Parliamentary approval, the practice has become established for each ISO to specify the purpose, the variables and data items required, and how the data can be used. ONS has found that excessively cautious regulations create three major problems:

   i. they lack the flexibility needed to operate effectively: they prevent reuse of data for other, previously unforeseen statistical purposes without a further ISO;

   ii. cautious drafting has sometimes made implementation of an ISO difficult because it cannot reflect the complexity of the operational systems on which the data are held. For example, the Disclosure of Social Security and Revenue Regulations were unable to be used in practice[2]: the wording of the Regulation placed limitations on the data that could be provided. This made it impossible for DWP to provide the data because of the way their systems were designed; and

---

[2] In this particular case an arbitrary "cut-off date" in relation to data subjects was inserted in the interests of "proportionality", and it was this cut-off date that proved the problem when it came to implementation. As a result, ONS had to resort to a work around based on s22 SRSA, to which DWP only agreed because Parliament had already expressed its will in the regulations.

iii. this approach is impractical where large-scale datasets with many attributes are involved (this can run to several thousands). Without new legislation proposed in this document, ONS assesses that it is likely that this cautious approach will continue.

d. The need to seek approval from Parliament before ONS accesses data makes it very hard for ONS to carry out the feasibility work required to develop the case needed to secure Parliamentary approval.

e. Once agreement has been reached that data sharing is justified and a new gateway is needed, the ISO is drawn up. Before it can come into effect, it must be approved by Parliament through the affirmative resolution procedure. Once before Parliament draft orders cannot be amended, if one point causes concern, the entire order falls. The Parliamentary procedures around affirmative resolutions add at least an additional six months to the overall time taken before data can be shared.

f. ISOs may not under s47(2) be used for ONS to acquire information to provide statistical services. These are defined in s22. This restriction is thought to have originated in a desire to ensure fair competition in securing survey work. However, it also prevents ONS from acquiring information for the purposes of statistical services in relation to public authorities.

Evidence for change

54. There are several benefits from providing easier, quicker, but safe access to identified data so that it can be used for statistics. These are:

a. Efficiency – maximising the benefits of administrative data held by Government by collecting data once and using it many times, and reducing the burden on businesses and other respondents;

b. Improving policy making decisions based on research and statistics by strengthening the evidence base for policies – enabling new statistics and fresh insights on social and economic change to be developed in a timely way so that they can contribute to public debate and inform policy makers early on; and

c. Improving the quality of statistics - access to a wider range of identified data will make statistics more relevant, more timely and more reliable, and can reduce some of the uncertainties around small but significant changes emerging from survey results.

55. This section sets out some examples where easier access to identified administrative data can help to improve the evidence base and accountability of policy and decision making.

<u>New Policy questions: new statistical outputs</u>

56. Some policy changes emerge over a period of time in response to broad, gradual societal change. Others arise very quickly, in response to sudden changes or specific events. Statistics must keep pace with these changes and provide evidence about the effectiveness of particular policy interventions. Increasingly, policy makers need to understand the impact of their interventions on different sectors of the economy or society. In these cases, new outputs are essential to inform wider debate about societal or economic changes.

Pensions

57. Currently, ONS does not have sufficient information about employer and employee contributions to pension schemes. Matching employee data from the PAYE system to employer records on the business register held by ONS would enable analysis by size and type of business, as well as estimates of the value of employee contributions to the National Employment Savings Trust (NEST).

Understanding the UK Economy at a Time of Change

58. National Accounts (GDP) are the primary indicator of the nation's wealth and of the health of the UK economy. Access to individual-level PAYE data could enable ONS to provide:

      i.   better quality estimates of the contributions of different industries to GDP and the income of people working in different industries;

      ii.   better data on the state of the economy in different parts of the country which would give policy makers more accurate information to develop local economic policies; and

      iii.   more rigorous quality assurance, based on individual data, improving the estimates for users and the transparency of production from source data to final estimate.

59. There would be additional benefits for ONS responding more quickly and effectively to new challenges, quickly developing new estimates to reflect the changing economy.

<u>Reduced respondent burden and reduction in survey costs</u>

60. ONS collects information through surveys, which people and businesses have already provided to Government for administrative purposes. Getting access to these data would allow ONS to reduce the burden placed on respondents to produce the vital macro-economic, population and social statistics that the UK needs to support policy making and inform debate.

Reducing the Burden on Business

61. Access to Corporation Tax and Income Tax data would contribute to on-going work to minimise respondent burden, reduce costs for businesses and for ONS. At present ONS business surveys require around 1.25 million responses from over 250,000 businesses each year. Responding to these surveys is estimated to take over one million hours and cost businesses over £22 million per year. Some of this information is already submitted to HMRC in Corporation Tax, Income Tax Self-Assessment and PAYE returns.

62. Giving ONS access to this information would enable the size and scope of surveys (e.g. Annual Business Survey and Monthly Wages and Salaries Survey) to be reduced, resulting in savings to businesses in the order of £4 million per year. The additional coverage and scope of administrative data would facilitate the production of better quality statistics and improve efficiency by reducing the need for business to provide ONS with information which they have already supplied to the Government for administrative purposes.

Improving Population Statistics

63. Good population statistics underpin resource allocation at the national and local level and are fundamental for policy formulation, decision-making, research and outcome monitoring. In addition, such statistics inform decisions on the allocation of regional aid and enable the UK to fulfil international obligations. Over the past three years the Office for National Statistics (the Beyond 2011 Programme) has been researching new approaches to counting the population. While this work has demonstrated the potential for the future production of population estimates, more work is needed. Following a careful assessment of the statistical research, the findings of an independent review of methods conducted by Professor Chris Skinner, public attitudes research and the responses to the public consultation, the National Statistician recommended that the Authority should make the best use of all sources, combining data from an online census in 2021 with administrative data and surveys. The increased use of administrative data will not only enhance statistics from the 2021 Census and improve statistics between censuses but offer a springboard to the greater use of administrative data and surveys in the future. Such an approach has the potential to improve the

accuracy, frequency and efficiency of existing statistics and the potential to provide new statistics for topics such as household income which could not be collected in a census because of concerns about data quality. The Government has welcomed the National Statistician's recommendation.

<u>Delivering efficiency and improving statistical quality</u>

64. Improving statistical quality has a direct impact on the quality of the evidence provided to policy makers, enabling decisions to be based on more timely, relevant and comprehensive information.

Improvements to the Business Register

65. The Inter-Departmental Business Register (IDBR) holds information on UK businesses and is widely used by the Government to provide information on the structure of the economy, for labour market statistics and to conduct surveys. The IDBR uses company registration data to help match VAT/PAYE records to identify which businesses to include in a business survey sample frame. However, company registrations do not indicate trading status or economic activity (many company registrations are made for non-trading purposes). The companies' registration system holds over 3m live companies, whereas the IDBR only contains around 1.4m of these. Of the remainder, it is possible that a proportion is actively trading but are under VAT/PAYE thresholds, and as a result are not included.

66. If ONS had access to Corporation Tax records this would enable identification of businesses that are actively trading, and would improve the coverage of small businesses. This could allow ONS to capture changes in the economy more quickly and provide more responsive analysis.

Labour Market Statistics

67. There are three key labour market series produced by ONS – workforce jobs; unemployment and employment measured through the Labour Force Survey (LFS); and the claimant count. Statistics on the number of jobs in the UK are collated from a quarterly business survey. Access to PAYE real time information from HMRC would potentially allow monthly rather than quarterly publication of these statistics and increase the accuracy of the figures by drawing on data from a much greater number of businesses. Policy makers looking at the labour market would have access to improved and more regular estimates of the number of jobs. Being able to consider these together with claimant count statistics and the sample survey-based estimates of employment from the LFS would allow policy makers to identify trends in the labour market with greater

confidence. Access to these administrative data would also enable an improved understanding of the characteristics of all three data series, which could facilitate quality improvements to the statistics in the future.

<u>Linking data for other public authorities</u>

68. As part of ONS's collaboration with the Administrative Data Research Centre (ADRC) for England, it will link data from public authorities, including data held by ONS, and provide disclosure controlled outputs to approved government researchers in safe-settings. This function is covered by existing provisions in the SRSA (s23 "promoting and assisting statistical research").

69. ONS will use these existing powers but have assessed it would be able to fulfil this function more effectively with a speedier alternative to the Parliamentary process for accessing information from public authorities.

<u>Options considered</u>

70. The proposal is to create powers that would enable identified data held by public authorities to be shared for the Authority's functions (which are primarily statistical). Earlier chapters set out the potential benefits that could result from streamlining access to identified data from administrative sources. This chapter states the options that have been considered. In doing this, we have been driven by the need to: prevent the misuse of data; ensure that the data acquired can be used only for statistical purposes; and, ensure that no identifiable information about individuals is unlawfully disclosed.

71. The following options were identified and appraised (full details are set out at Annex E):

- Option 1 - No change to the existing arrangements;
- Option 2 - Remove restriction in SRSA s45 (5) on personal HMRC information;
- Option 3 - Permissive power for the Authority; and
- Option 4 - Broad power for all public authorities to share identified data with each other for research and statistics. At the Open Policy Making Plenary meeting on 22 October 2014, members of Civil Society, present in person and by proxy, made it clear that Option 4 would be a step too far and would be a red line for them. It will therefore not be taken forward.

Existing Safeguards That Will Continue

Legal safeguards specific to The Authority[3]

72. Disclosure of information held by the Authority is restricted by law: s.39 SRSA contains a criminal penalty for unlawful disclosure. In addition, the Authority and ONS are subject to the Data Protection Act, the law of confidence, and the Human Rights Act 1998, none of which would be amended by the Authority's proposals. Any breach of the Data Protection Act could result in a fine being levied by the Information Commissioner's Office. In addition, data disclosed to ONS are subject to specific terms and conditions agreed between the data owner and ONS. Similarly, unlawful disclosure of information covered by the Social Security Administration Act or the Commissioners for Revenue and Customs Act 2005 would constitute a criminal offence under these acts. Researchers approved under s.39 (4)(i) of the SRSA would also be subject to contractual constraints and penalties in accordance with arrangements set out in data access agreements made by the Authority; ONS employees are subject to disciplinary procedures under their contract of employment.

Governance

73. ONS is independent of ministers and, as the executive office of the Authority, operates at arm's length from government. Governance of the Authority is set out elsewhere in this document. The SRSA sets the Authority the objective of promoting and safeguarding the production and publication of official statistics which serve the public good. The Authority and ONS are unable to exercise any functions outside the SRSA.

Policy safeguards

74. The Authority and ONS must comply with the Government's Security Policy Framework. This provides the basis for assessing and managing risks and protecting key information assets. As a result ONS must uphold standards for information assurance, data security and risk management including those promulgated by CESG (the National Technical Authority for Information Assurance) and the International Organisation for Standardisation (IOS). Such standards cover: data transfer, systems, procurement, reporting and training.

---

[3] In general, Authority would remain subject to the Data Protection Act, the law of confidence, and the Human Rights Act 1998, in respect of the information it received. A serious breach of the data protection principles also attracts liability for monetary penalties levied by the Information Commissioner's Office. Data disclosed to ONS is subject to specific terms and conditions agreed between the data owner and ONS. Researchers approved under s39(4)(i) of the SRSA will also be subject to contractual constraints and penalties under data access agreements under which the Authority discloses information to them; ONS employees are also subject disciplinary procedures under their contract of employment.

75. Where relevant, other processes/procedures may include the completion of the appropriate type of privacy impact assessment and/or an independent review of security arrangements (e.g. those undertaken for the 2001 and 2011 Censuses). In addition, ONS would need to comply with specific departmental requirements/conditions including clearance or approval from bodies such as the Data Access Ethics Committee in DWP and the Data Management Advisory Panel in DfE; including the ability to fully inspect the facilities, and audit data handling processes and procedures.

76. All staff must sign the ONS Confidentiality Declaration to confirm that they understand their obligations to keep information safe and secure and the penalties associated with any infringement of ONS statutory and other related obligations.

77. The ONS Information Charter[4] explains how ONS carries out its responsibilities for handling personal information (in addition there is a 'Respondent Charter for Business Surveys' and a 'Respondent Charter for Households and Individual Surveys). Easier access to administrative data, and being able to match and link information for statistical purposes will help ONS to meet specific pledges in their Information Charter only (e.g. to ask for what is needed).

78. Ethical requirements are contained in the Code of Practice for Official Statistics, especially those covering Integrity, Confidentiality and the Protocol for Use of administrative sources for statistical purposes.

79. Any direct collection of data for testing or evaluation purposes complies with the principles set out in the Code of Practice for Official Statistics. The Code contains principles and practices that are intended to ensure that: the range of official statistics meets the needs of users; that the statistics are produced, managed and disseminated to high standards; and that the statistics are well explained.

Security

80. Government has set out, in its HMG Security Policy Framework[5], the standards, best practice guidelines and approaches that are required to protect UK Government assets. This sets the minimum obligations for the Authority. Personal accountability for data is ensured by the requirement, under this Framework, to appoint and train, for each data asset, an Information Asset Owner (IAO), who is responsible for it. ONS has produced a handbook[6] for the

---

[4] http://www.ons.gov.uk/ons/about-ons/business-transparency/information-charter/index.html
[5] Version 11 – October 2013
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/299556/HMG_Security_Policy_Framework_v11.0_doc.pdf
[6] The Information Asset Handbook (ONS v1.8 May 2014)

use of IAO. Data are transferred in accordance with the Security Policy Framework. When required CESG-approved encrypted media are used with encryption passwords and/or tokens controlled by either ONS Security Managers or the Security Managers in the owning Department.

81. ONS has published examples of their approach to safeguarding data[7]. In accordance with Government requirements; when working with data, ONS imposes the following controls:

- Physical security - access to ONS buildings is controlled and monitored;
- Personnel security- all personnel are subject to security checks to the level required for their role; and
- Procedural security – all data acquisition, import and export processes are subject to strict procedural controls, in many cases incorporating separation of duties.

82. ONS recognises the security risks of handling identifiable data and has taken some specific measures when linking and matching across disparate datasets. For example, data anonymisation processes were developed for the Beyond 2011 Programme. ONS has implemented a range of processes to ensure that appropriate levels of anonymity and privacy are maintained where appropriate.

83. Data export and publication are carried out in accordance with the SRSA and Code of Practice for Official Statistics whereby no personal information about an individual is disclosed in any statistical output. All outputs from ONS research are subject to Statistical Disclosure Control (SDC), that is methods designed to protect individuals, households and businesses (and their attributes) from identification in any published tables or other statistical outputs.

Procedures

84. The Authority has set out the process that is currently undertaken when negotiating and agreeing the acquisition of data under the current SRSA regulations (see Annex F). The purpose of these procedures is to determine that:

- the information is of sufficient quality;
- the information is actually required;
- the proposed data share complies with existing legislation including the DPA and the Human Rights Act; and
- privacy risks and issues have been addressed appropriately;

---

[7] http://www.ons.gov.uk/ons/about-ons/who-ons-are/programmes-and-projects/beyond-2011/reports-and-publications/beyond-2011-safeguarding-data-for-research-our-policy--m10-.pdf

- security requirements and standards have been met and account has been taken of its impact on the business and statistical outputs of ONS.

85. The terms used in Annex F are explained in a document entitled "Stepping Stones", which provides guidance for members of the Government Statistical Service to use when considering data sharing applications for statistical or analytical purposes[8].

Accreditation process

86. The Authority already has the power (s.39) to disclose personal information that it holds to an approved researcher for statistical research. It publishes the criteria applied to secure accreditation as an approved researcher as well as the measures taken to assess the suitability of individual research projects[9].

New oversight safeguards options as alternatives to the Parliamentary Process

87. Under SRSA s.47, ISOs are approved by affirmative resolution in both Houses of Parliament. Affirmative resolution fulfils two functions:

- legal authority for the data share; and
- independent scrutiny of the proposal to ensure that the business case is robust and the conclusions justified.

88. ONS has found that affirmative resolution process adds approximately six months to the time it takes to get a data share, therefore new options are proposed as alternatives to the Parliamentary process. Any alternative approach to scrutiny and decision making must have the same, or greater, rigour; ONS does not wish to limit scrutiny in any way. However, in order to achieve increased efficiency and flexibility to inform timely policy decisions ONS is keen to ensure that legal approval for any proposed data share can be made quickly and easily.

The current approach: Information Sharing Orders and Affirmative Resolution in Parliament

89. Many steps precede the Parliamentary process to ensure that a proposal is appropriate and lawful. There are extensive discussions between ONS and the data owner to identify and substantiate the requirements for access to the data. ONS must conduct a full legal review, establish the statistical and business case and carry out a privacy impact assessment. Governance arrangements vary between data owners: in some cases the proposed data share must be reviewed

---

[8] http://www.ons.gov.uk/ons/guide-method/best-practice/gss-best-practice/stepping-stones-to-data-sharing-for-statistical-purposes/index.html
[9] http://www.ons.gov.uk/ons/about-ons/business-transparency/freedom-of-information/what-can-i-request/approved-researcher-accreditation.html

by a departmental ethics committee or equivalent, in others it must be reviewed at Board level. These steps provide internal scrutiny and ensure that, from the perspective of the data owner, the share is appropriate, legal, proportionate and ethical. This work is essential and should continue to be part of any data sharing process.

90. Once Officials, lawyers and Ministers agree that an ISO is appropriate, the Minister for the Cabinet Office (MCO) writes round to all Ministerial colleagues to secure their agreement to the Order. If Ministers are content and satisfied that the conditions at s.47(9) are met (including the public interest test), the ISO is reviewed by the Joint Committee on Statutory Instruments (JCSI) to assure that it is properly drafted and within the legal powers conferred by s.47, and the Secondary Legislation Scrutiny Committee, which considers the proposal to identify whether it:

- raises issues around legal, political or public policy; or
- is inappropriate because circumstances have changed since the relevant primary legislation was passed; or
- inappropriately implements EU legislation; or
- imperfectly achieves its policy objectives.

91. Once the regulations have been scrutinised and endorsed they are laid for consideration by relevant committees in the House of Commons and the House of Lords. A short debate follows, informed by any issues raised by the Scrutiny Committees. There is no scope to amend the Order; it stands or falls as laid. Affirmative Resolution provides the opportunity for scrutiny by Ministers before the data owner gets the legal permission to share data.

92. Hansard shows the depth and nature of the debates for the five ISOs that have been laid. The number of attendees at the debates has ranged from two to fifteen; the length of debates is usually 15-30 minutes; the longest debate lasted 40 minutes (this was the first ISO), the shortest debate lasted one minute. The debates rarely consider the details of the data share, but instead discuss wider issues related to it (for example, the importance of the Census, or data security). Assuming the Order is approved by both Houses, it is signed by the Minister for the Cabinet Office and other relevant Ministers and becomes law.

93. There have been strong representations from some elements of Civil Society that the Parliamentary process should be sacrosanct.

Possible alternatives to Affirmative Resolution

94. A key principle that has underpinned the consideration of alternatives to Parliamentary scrutiny is the need to ensure that decisions are made at an appropriate level, by a person or body which can be held to account.

95. The place for independent external scrutiny of data sharing proposals has also been considered. Independent scrutiny provides support to those making decisions about complex or unfamiliar issues, and assures the public that the proposal has been considered from an external perspective. It should be transparent, with decisions and advice made public. It should be rigorous, giving detailed, expert consideration to each proposal, and provide assurance that the proposal:

- is in the public interest;
- is lawful;
- supports a valid statistical purpose; and,
- appropriately reflects practical issues such as compliance with government security policies and standards.

96. Three options have been identified (details are provided in Annex E):

- Option 1: ISOs approved by Affirmative Resolution (Do Nothing)
- Option 2: Decision by Minister
- Option 3: Involvement of an Independent Ethics and/or Approvals Body in the decision making process. There are four further variants of this option.

Next Steps

- Further consideration and development of alternative options for Affirmative Resolution
- Engagement with the Devolved Administrations to assess their statistical institutes' requirements for identified data.

C) HMRC Strand – Sharing general, aggregated and de-identified data for public benefit

The situation

97. HMRC is a statutory body, created by the Commissioners for Revenue and Customs Act 2005. This imposes a duty of confidentiality on HMRC officials, which applies to all information that HMRC holds in connection with its functions. HMRC takes this duty of confidentiality seriously. A criminal sanction protects against the unlawful disclosure of, information that identifies a person or through which their identity can be deduced (called "identifying information" in this paper). HMRC may share information only in limited circumstances set out in legislation, in particular:

- For the purposes of HMRC's functions; or
- With the consent of each subject of the information; or
- Through specific legislative gateways ('statutory enactments')

98. Once we have a valid legal basis enabling disclosure, HMRC must consider any data sharing options or proposals, i.e. the need to ensure compliance with the Data Protection Act and Human Rights Act, alongside practical elements - resource implications etc.

99. HMRC holds sensitive information and it accepts that it is right for there to be a strong focus on any information sharing proposals. However HMRC holds a spectrum of information ranging from non-identifying through to identifying information that is extremely sensitive in nature. It is therefore arguable that the current protections offer more protection to, for example, non-identifying information, than is needed and that a more tailored approach could be taken, accounting for sensitivity and risk, with appropriate safeguards to ensure that confidentiality is not compromised.

100. HMRC identified three specific information types that it considered to be at the lower end of this spectrum and last summer consulted on specific proposals for the wider sharing of these information types, including a proposal to share general, aggregate and de-identified data for purposes wider than HMRC's functions to generate public benefits.

General and Aggregate information

101. General information is information that is not, nor ever has been, identifying information, for example, information on policies and processes.

102.   Aggregate information is grouped information, summarising the characteristics of a set of data. This is potentially more disclosive than general information, but still generally low risk within the spectrum of information types that we hold because it is not disclosed on an individual-level basis. Where HMRC is currently able to disclose this type of information, it does so using safeguards that are appropriate to the data type. This includes employing strict security and information management processes, and robust statistical disclosure policies. Permissive gateways mean that disclosure is not mandatory and the criminal sanction protects against unlawful disclosure of identifying information (which could occur if, for example, the aggregation was at too granular a level).

103.   If HMRC could share this information more widely by way of a broad gateway enabling disclosure of aggregated data for the purpose of delivering public benefits, HMRC could contribute to the more efficient and effective delivery of services and benefits beyond HMRC's functions, for the benefit of UKplc.

De-identified data

104.   De-identified data cannot directly identify an individual, and so does not amount to personal data under the first limb of the definition of Personal Data under the DPA. This data could nonetheless potentially amount to personal data under the second limb of the definition if individual to which it relates could be identified from the combination of that data with other data held or likely to be held by the data controller.

105.   HMRC currently provide access, on this basis, to this type of data for research purposes. However, those research projects must be able to demonstrate a benefit to HMRC's functions, limiting the potential to deliver research for public benefits beyond HMRC's functions.

106.   Recognising the greater risk of a customer's identity being deduced than in the case of aggregate data, the following safeguards are currently in place and will remain unchanged under the proposal:

- A secure and controlled environment provided by HMRC's Datalab[10], which has been operating successfully for over 3 years;

- Only projects with a valid research purpose and from trusted organisations are allowed. HMRC expects publication of the findings from the research;

---

[10] http://www.hmrc.gov.uk/datalab/about.htm

- Users undergo a rigorous accreditation process and need to sign an agreement with HMRC on the use of the information;

- Datasets are de-identified, and statistical disclosure controls are carried out on any research outputs before they leave the Datalab;

- Researchers are subject to the same confidentiality provisions as HMRC staff, including the criminal sanction; and

- The environment and processes are consistent with the recommendations in the recent Administrative Data Taskforce Report for the safe sharing of data for research and statistical purposes.

107.   The proposal seeks to provide a legal gateway which will allow research to be undertaken for wider public benefit and not just, as currently, for HMRC's functions.

Evidence for and against change

108.   Left as the status quo, HMRC will, as now, be approached with requests to disclose information, which will be considered on a case-by-case basis. If a valid legal basis is available that could allow disclosure, HMRC will need to consider any data sharing options or proposals, i.e. the need to ensure compliance with the Data Protection Act and Human Rights Act, alongside practical elements - resource implications etc. However if a valid legal basis is not available, this has to be provided for before disclosure can be made. A legislative vehicle needs to be found and the process of creating a statutory gateway can typically take up to two years.

109.   The proposal would enable a broad gateway to be implemented by reference to information type (i.e. general and aggregated data, de-identified individual level data), where this would lead to 'public benefit'. Having this gateway in place would allow HMRC to contribute to wider government initiatives than it currently can and for purposes beyond HMRC's own functions. In particular, a broader gateway could improve the evidence base for policy-making and promote knowledge sharing between research organisations and the public sector.

110.   The absence of a legal gateway can frustrate wider policy formulation and development and addressing these data needs by the usual way of a new legal gateway on a case-by-case basis is time-consuming and resource-intensive. HMRC accepts that identifying information is particularly sensitive and should be subject to rigorous and on-going scrutiny and critical assessment. However if less sensitive data types (with safeguards as appropriate) were available under a broad gateway, not only will HMRC be contributing more effectively to wider initiatives with a view of delivering public benefits on a broader scale, but

government departments could be encouraged to seek less sensitive data by way of this existing gateway (if implemented) rather than the default of seeking a new gateway for potentially identifying information.

111.   Listed below are some examples of approaches for HMRC's information that have either had to be turned down or have had to be substantially modified:

- An approach was made to HMRC to supply anonymised data to help the Chief Medical Officer for Wales to carry out research into the factors underlying excess winter mortality. Considerable work was undertaken to identify whether and how the information could be disclosed. Obtaining customers' consent to disclose anonymised data was not practical, there was no link to HMRC's functions and legislating a specific gateway would have been too time consuming and would, in any case, have been too late to inform the research. The conclusion was therefore reached that it would not be possible for HMRC to supply the requested information.

- BIS applied to use the Datalab to produce tables of profits turnover, counts of total and loss making companies by turnover size and industry sectors using Corporation Tax. The aim of their project was to:

  o   investigate the relationship between company profitability and deposit holdings for the non-financial corporate sector by company size and industry sector; and

  o   quantify the issue of companies only able to pay interest on debts with no ability to invest and grow their business.

  This was rejected by The Datalab Committee because assessing companies' productivity did not fall within HMRC's functions.

- The LSE were hoping to match HMRC data to the Annual Business Inquiry from the ONS to inform a paper on the determinants of outsourcing of business services, a sector that plays a vital role in the U.K. economy (Globalisation, Managerial Complexity, and Service Outsourcing). This was rejected as it did not fall within HMRC's functions.

- The Bank of England was unable to conduct research to investigate the relationship between the prices charged by individual firms and total sales, their costs and other characteristics. They wanted to do this by analysing income (and productivity) profile of self-employed individuals and matching data from Consumer Price Index to HMRC PAYE data (to get information on wages) and VAT returns (to get sales, inputs and

value added) within the Datalab. This approach was unsuccessful as the project was beyond HMRC's functions.

112.   In addition, responses to the public consultation in Summer 2013 were supportive of this proposal, as long as the safeguards were sufficient to protect confidentiality (particularly in case of de-identified data). As noted above, HMRC already makes this type of information available, with strict safeguards, for research that links to HMRC's functions and has done so successfully for over three years. However HMRC understands the concerns raised and, as previously noted, will explore other proposed safeguards as part of the open policy making process.

Options identified and appraised

113.   The proposals for wider sharing of general, aggregate and anonymised data set out in this proposal and our earlier consultation document were informed by HMRC's experience of disclosing these data types, where it is currently legally able to do so. HMRC was able to offer up safeguards that are currently applied successfully, while seeking views on these.

114.   It is proposed that the purpose of the legal gateway should be framed in terms of delivering public benefit. Public benefit would be judged by an approvals committee within HMRC, which might include external representation. The argument that such a body should be within HMRC rather than an external body, is that it is de-identified data that is being provided. This is similar to the approach adopted by UK Statistics Authority in disclosing information in its Virtual Microdata Laboratory to accredited researchers, as empowered under s39(5) of the SRSA. This is informed by the experience of earlier unsuccessful requests for HMRC data which had to be turned down because, in order to provide the data, we specifically required a link to HMRC's functions. An alternative approach would be to specify a purpose 'beyond HMRC's functions', but this would result in a far broader gateway than a provision tied to 'public benefit'; in addition, 'public benefit' is considered more in line with the public's expectation of what Government departments' policy and initiatives should be framed around. However HMRC is aware that there needs to be clarity on what exactly is meant by public benefit and welcomes the open policy making process as a means by which this can be explored. For example there is a formulation of "public good" in the Statistics and Registration Services Act (s7(2)) 'public benefit includes in particular (a) informing the public about social and economic matters, and (b) assisting in the development and evaluation of public policy'.

<u>Proposed Method for Delivering the Recommended Approach</u>

115.　HMRC proposes that the legislation should provide the necessary structural framework for a permissive (not mandatory) legal gateway for each of the information types (i.e. for general and aggregate information and for de-identified individual level information) together with the purpose, alongside the main safeguard of a criminal sanction protecting against unlawful disclosure (of information that identifies a person or through which their identity could be deduced). However HMRC considers that in order for the gateway to have sufficient flexibility for the future, that there will need to be elements that are maintained outside of legislation, for example through a policy statement and/or a code of practice which could be provided for in statute. HMRC would be looking to develop this aspect, picking up on similar elements being developed by the new Administrative Data Research Network.

116.　By way of illustration, a policy statement could be used to set out the statistical disclosure tests that HMRC applies to aggregate information, with the aim of ensuring that it will not be possible to deduce information about identifiable individual persons from aggregate information; the criteria used to assess public benefit; governance of Datalab research requests[11]; accreditation/vetting processes; publication of the identities of those approved as accredited researchers; publication of the subject matter of research that has been approved; and the requirement for a plain English summary of the outcome of the research to be published.

117.　This is a different proposal to the Trusted Third Party strand of the data sharing proposals. This is because in these proposals there is a statutory bar within the Commissioners for Revenue and Customs Act 2005 (CRCA) which needs to be amended in order for HMRC to make data available to others in HMRC's own safe setting.

118.　HMRC currently asks other government departments to cover HMRC's costs in providing data to them and would expect that other government departments would account for this element, when determining the costs and benefits of their policies.

<u>Next steps</u>

119.　Development of alternative options on how HMRC could use its data more effectively by, for example, delivering benefits beyond HMRC's own functions, in balance with protecting confidentiality.

---

[11] For example HMRC is open to considering the establishment of an HMRC-led advisory committee with independent representation to consider applications to access HMRC data for the purpose of public benefit.

2 - Fraud, Error and Debt

The situation

120.   Fraud and Error, as a cost to the whole UK economy, stands at **approximately £73bn with approximately £20.3bn being attributable to the public sector**. [12] Whilst this is an indicative figure it may not be completely accurate as there are unquantifiable considerations, such as the activity of the shadow economy, which will always require a degree of educated assumption. The work carried out by the Fraud, Error and Debt team within the Cabinet Office estimate that this cost could be within the range of **£38bn - £67bn**. This represents a range of **2.62% - 5.03% of GDP**. This has been based on the data from the Annual Fraud Indicator and comparator data from European countries and the US as well as current thinking on issues such as the shadow economy.

121.   During discussions, government estimations of fraud and error were presented setting out an estimated £22bn known fraud loss and a loss of approximately £14.5bn in relation to error. These were based on the Annual Fraud Indicator (mentioned above) and are therefore subject to the same issues as set out above. They are also complicated by factors such as the definition of error not being uniform, and therefore rightly, were questioned by the group. Our proposals seek to gather evidence that would give a greater understanding of the full costs of Fraud and Error.

122.   If a Government department's recorded Fraud and Error is above **1% of their Departmental Expenditure Limit** the National Audit Office are empowered to investigate the activity of the department. It is agreed that for some departments and public authorities this is an challenging threshold (with anecdotal evidence of private sector tolerance being set much higher than this at between 3-5% depending on the business); however, it is agreed that, whilst this is the trigger point for external investigation, efforts should be made to reduce losses through fraud and error to below that threshold; it is not used to indicate an 'acceptable' level of losses.

123.   The recent National Audit Office report (*Managing Debt Owed to Central Government*, February 2014) reaffirmed the importance of reducing debt owed to Government as part of good financial management. The Cabinet Office and HM Treasury are working jointly to further strengthen financial processes across departments through a range of measures.

---

[12] The most recent Annual Fraud Indicator (published March 2012 and found at: https://www.gov.uk/government/publications/annual-fraud-indicator) document sets the cost of Fraud and Error to the UK Economy as a whole as £73bn and provides a useful breakdown of this by sector.

124.   It is agreed that the current size of debt owed to the Government is £22bn. Of the overall debt that is owed to Government, 88% was owed either to HMRC or DWP, 10% was owed to the Ministry of Justice and 2% owed to other departments.

125.   It is agreed that this £22bn figure does not represent the size of *collectible* debt owed to Government. We are agreed that the potential sum is still likely to be significant enough to merit intervention to seek to increase the recovery of collectible debt, but this may require further investigation.

126.   The Government's two key aims are to provide better support to citizens to help them manage their debt as well as to increase the amount of debt collected by the Exchequer.

Evidence for and against change

127.   Wider use of data sharing could improve the prevention, detection and investigation of **Fraud and Error** by:

   a. aiding better targeting and risk-profiling of potentially fraudulent individuals;
   b. saving taxpayer's money by streamlining processes; and
   c. increasing the ability for Government to act more quickly on fraud and error by simplifying the legislative landscape.

128.   There are clear calls to increase the effectiveness and/or the efficiency of current data sharing from across the public sector and some private sector organisations. These are based on a reported lack of flexibility (the difficulty in adapting to changing circumstances in a timely fashion given current legislative processes), the complexity of navigating the current legislative landscape and the time taken to create new data sharing relationships against the opportunity cost of not reducing that time (i.e. the cost of vulnerability being exploited whilst time is taken to reduce that vulnerability). A working example of the issues faced by the Charities Commission when they set up a successful data-matching pilot is provided in Annex H.

129.   Simplification of current processes, either through legislation, or through increasing public sector skills, knowledge and general capability in data sharing matters, would be beneficial to government. A case study of the arrangements in place between Department for Work and Pensions and HM Revenue and Customs, which has simplified their data sharing through the use of a broader gateway, is provided in Annex I.

130. For **Fraud and Error**, a number of challenges and clarifications are needed to fully understand the evidence base and therefore provide a robust basis from which to assess whether further intervention is required:

   a. What **barriers** frustrate data sharing for Fraud and Error, what are the **incentives** that drive data sharing?
   b. What existing gateways aren't being used? What is the level of public official awareness of what can and can't be shared?
   c. What is the **public attitude** to data sharing for Fraud and Error?
   d. What are the **costs and benefits** of improved data sharing for Fraud and Error, can they be fully quantified (costs and benefits as defined in their broadest sense: privacy, financial etc)?
   e. What is the comparative value of **different approaches** (data analysis as opposed to validating data or validation and analysis in combination; case by case validation and analysis as opposed to bulk data validation where it is necessary and proportionate)?
   f. How do we **strike the right balance** between efficiency and effectiveness on one hand, and privacy on the other?

131. Wider use of data sharing could increase **debt** recovery by:

   a. ensuring the right information is available to the right department at the right time to identify debtors;
   b. making effective interventions earlier to prevent debt from accruing;
   c. making the process fairer, through a better understanding of the circumstances of the debtor;
   d. informing strategic decision making about debt management; and
   e. making debt collection more efficient, supporting work for creating a single point of access for the debtor and Government to engage. This more efficient process is also expected to be more effective, by making repayment clearer and easier for the debtor, although there are potentially issues for both public perceptions, definition of terms, and incentives that will need to be explored and resolved.

132. For **Debt** there are some key definitional issues that need to be agreed across government before exploration or clarification of evidence can take place. The OPM group challenged what Government was talking about when it talked about debt and how Government segmented this as different approaches may prove fruitful in different situations.

133. Once the definitional matters have been agreed the full benefits need to be explored and quantified in order to fully understand the value of intervention and its impact on collectible debt.

134.   In order to better understand the value of intervention, the OPM group agreed more clarity needs to be given on:

    a.   collective terminology - different public authorities define debt in different ways and have different sets of terminology surrounding their activities to manage debt and there is a wider definitional issue of what type of debt we are talking about that will determine best approaches to collecting (for example individual debt as opposed to company debt);

    b.   the drivers for change;

    c.   an understanding of what the size of *collectible* debt (as opposed to the overall debt balance) is across Government;

    d.   the impact that data sharing would have on raising the amount of collectible debt and helping citizens to manage their debt;

    e.   constraints needed to control disproportionate sharing of data; and

    f.   Which areas of debt to best target here.

135.   Given the varied definitional issues that need to be tackled with Debt, it was felt that the group could not progress further into proposals. However, it is likely that a number of the issues that the proposals on Fraud and Error are seeking to address could be considered for Debt. As such any methodology should seek to derive evidence, where possible, for the impact of intervention on debt.

Barriers

136.   There are a number of barriers that may frustrate the sharing of data, including the complex legal landscape and the resulting risk averse cultures. Two additional key drivers related specifically to fraud, error and debt are:

    a.   **Resources** - Sharing data requires resource. Organisations that hold a lot of data are often responsible for major functions of government (the administration of the benefits system for example) and have to make hard choices about how best to deploy their resources. At a more personal level, if data sharing is not somebody's 'day job' it is unlikely to be a priority for them to respond to requests to do so; and

    b.   **Financial incentive** - This is an issue for debt in the main, but is listed here. Organisations see debt as a revenue stream (albeit a net loss once collection costs and uncollectable debt is taken into account), if they collect that debt the money recovered offsets a proportion of the costs of recovery. Consequently any attempt to improve the central

management of debt without appropriate financial incentives in place will need to bear this in mind - particularly when it comes to distributing monies collected centrally, if this was a part of the solution.

137. It is clear that without a full understanding of the drivers and incentives for data sharing, then a proposal may be put in place that may not then be fully effective. At the same time without testing out some new approaches, it is unlikely that a full understanding of how the incentives operate will be reached.

Costs, Benefits and the Comparative Value of Different Approaches

138. A lot of the proposed benefits of data sharing in this area are derived from assumptions relating to isolated case studies that have proved successful. For a robust decision these benefits need to be clearly articulated and underpinned with evidence that can be scaled up to the point where it is effective without losing sight of the costs of intervention. Testing out potential interventions would improve our understanding of what the benefits would be and how easily they could be scaled up.

139. The costs of improved data sharing here also needs to seek to quantify, as far as possible the individual cost of intervention on top of the financial and resource implications of different approaches. Whilst some of the costs may be clearer the social impact may not be fully understood without testing out some of the proposals in a smaller environment.

Getting the Balance Right: Privacy v Effectiveness and Efficiency

140. In order to ensure that a solution is effective, it needs to take into account the financial costs and benefits as well as minimises intrusion on individual privacy, particularly of individuals who would not be of interest in relation to fraud. To achieve this balance, evidence would need to be gathered about how effective any proposed solution would be in minimising cost in particular and ensuring that principles of necessity, proportionality and transparency are applied appropriately. We acknowledge the benefit to citizens of validating/verifying information using data already held by government rather than alternative processes which may be more intrusive and inconvenient.

141. The Group proposed that Privacy Impact Assessment principles be embedded throughout the work on fraud, error and debt. The group recognised the value of adhering to the current best practice in order to ensure that proposals are balanced as well as support greater transparency and accountability along with the other safeguard principles adopted earlier in this paper.

142.   In order to deliver robust, evidence based assessments of potential options we recommend further evidence gathering:

- a set of **surveys** to gain better insight into the public understanding of what the Government currently does with data to tackle fraud, error and debt and the tolerance that the public has for further sharing - This, alongside a more in-depth deliberative research study, will help us to determine potential tolerance levels for intervention in this field;

- the development of a set of **case studies** where data sharing to reduce instances of fraud, error and debt has and has not been successful in achieving its objective, with a clear understanding of what has incentivised good sharing and what has driven poor sharing - building up a more in-depth view of the costs and benefits of intervention as well as lessons to draw from previous attempts - this includes the gathering of evidence from currently planned or running proof of concept and pilot projects; and

- a set of **trials and pilots** for ways of improving data use in order to understand the comparative value of different types of intervention and deliver even greater understanding of the **impacts of intervention**. Two pilots have been suggested so far: one that tests an approach that seeks to filter out low-risk citizens so that more resource can be focussed on those with a higher risk of committing fraud; and a second that allows citizens to correct erroneous information held about them when decisions are made.

143.   Whilst this evidence is gathered we recommend a series of assessments of the evidence for intervention based on full costs and benefits analysis that takes into account effectiveness, efficiency and the potential privacy impacts:

- We recommend that these assessments are carried out by a panel (similar to the Cabinet Office ID Assurance Privacy and Consumer Advisory Group), constituted of Central and Local Government representatives, representatives from Civil Society organisations;

- The panel ought to be jointly chaired by the Cabinet Office and a Civil Society representative with a member of the Information Commissioner's Office attending in an advisory capacity;

- We recommend that any assessment and evidence contributed to it is made publicly available in order to aid transparency; and

- We recommend that the results of the assessment should form the basis for the Government's decision to take forward proposed intervention.

Method of delivery

144. We propose a number of evidence gathering projects that will help to provide greater insight into the situation and provide a robust basis for any decisions on solutions. These include a mixture of surveys, studies and pilots. At the end of this programme of work, Government should have the evidence it requires to either continue to implement improved data sharing here, or seek alternative solutions. A summary of the objective and delivery approach for each of the recommended actions is set out at Annex J.

145. We would seek to ensure that throughout the process there is inclusive, continuous and transparent governance of this programme of work, through a panel constituted of representatives from central and local Government, Civil Society and the Information Commissioner's Office with a joint Cabinet Office and Civil Society chair.

146. The panel would have oversight of the set-up, gathering and assessment of evidence, being responsible for ensuring that at every stage an unbiased and fair gathering of evidence is carried out. They will be responsible for the assessment of that evidence and will ensure that the results of any assessment are transparent and available to the public for scrutiny.

147. At the end of the programme it is proposed that the panel will make a full assessment of the evidence gathered and recommend the agreed way forwards.

Next Steps

- Further work to define debt and how it is understood and segmented across government with a view to determining whether it remains within scope of this work.
    - Civil society groups were concerned about the potential changes to debt management practices brought in by data sharing and it was agreed that further discussion is needed.

- Development of detailed implementation plans for pilots, surveys and delivery of case studies.

3 — Tailored Public Services

<u>The situation</u>

148.   For Britain to lead the world in transformative public service design and delivery in the context of increasingly reduced finances, public agencies will need to work ever closer together to identify new ways to support citizens. Legal restrictions around data sharing between public agencies are increasingly frequently cited as a critical barrier to the design and delivery of public services in new ways. The current legal context is highly complex, and multiple barriers prevent data sharing across and within public agencies. These hinder the ability of the government to protect the most vulnerable and to achieve improved outcomes for citizens. The results of barriers to data sharing include:

- public agencies are not able to share data to identify accurately which citizens are eligible for a particular service or benefit and therefore to ensure that the right people receive it;

- overlap, contradiction and gaps between services provided to an individual; citizens receive disjointed and fragmented services from a range of agencies as their needs change during their lifetime;

- reactive service delivery which means citizens often receive support too late; and

- public agencies cannot free up time for citizens by exchanging information internally so citizens are required to repeat their information multiple times and spend significant energy and time meeting bureaucratic requirements.

149.   To address these issues and to ensure that public services are delivered in a manner which is person focused the following policy has been developed through a process of iteration and testing with multiple stakeholder groups. The objective of strand is to maximise the benefit of data already held by public bodies to deliver public services tailored to individual needs.

<u>Recommendation</u>

150.   'A permissive power for *defined public agencies* to share data with *defined public agencies* for the purposes of improving the delivery or targeting of public services where it supports the achievement of the defined objective and where, in each share, a *majority of the individuals* whose data is shared are offered, as a

result of the share, an intervention[13] which intends to support them to improve their *health, education or employment outcomes*.'

Phrases in italics indicate an area for further discussion (see below).

151.   The term 'public bodies' in the proposed power is not a pool of agencies which can share data, but rather the ability for specified public agencies to be able to share data with other specified public agencies/ specified departments in a local authority, in a defined direction, determined by policy need. Current public agencies recommended for inclusion are:

- Central government departments;

- Agencies of central government departments;

- Local authorities;

- Local health care providers and commissioners. Use of this power by local health providers, or by other parts of the public sector seeking access to health data, would be conditional on involvement of an appropriate expert panel, such as arrangements similar to S251 of the NHS Act 2006 and subject to oversight by the Independent Information Governance Oversight Panel; and

- Devolved administrations.

152.   With the increasing fragmentation of public service delivery across the public, private and third sectors, there will increasingly be challenges to the efficacy of a power which limits data sharing to public agencies. Should this include organisations functioning across the public sector, such as private, not-for-profit and voluntary sector contractors, and what additional safeguards or restrictions might this require? We recognise that this will have impact on this policy and will need further work to explore the different options and associated implications.

153.   The intent of the term 'Majority of individuals' is to ensure protection of individuals' personal data through guaranteeing that there is a strong link between peoples' data being shared and a direct benefit being offered to them. However, the following challenges to this approach are evident:

---

[13] The intended meaning is any act which aims to support the citizen, so that could be an invitation for an assessment of need by a local authority, receipt of a financial benefit, a reduction in bills, offer of a service etc.

- There are difficulties with both the use of the word "majority" and the alternative of trying to specify in law a certain proportion of people who will receive an intervention. It may be better to remove this entirely and rely on necessity and proportionality required by the DPA and HRA. Data sharing can be justified on the basis of supporting decisions taken in respect of particular individuals, for example to identify whether a person is eligible for/ most in need of a targeted service or benefit, rather than on a specified proportion of people receiving a certain intervention. Further work will need to explore the implications of including the current wording or options for alternative wording, to ensure the policy intent is achieved.

- An additional challenge is that it might not be known before doing the data match whether a majority of individuals will benefit and therefore whether the use of the power is lawful. In this case a 'trial data match' should be undertaken and the requirement for a majority of individuals to benefit directly is not applicable; the requirement is instead that there needs to be a reasonable theoretical explanation for why the share is expected to result in an intervention being offered to the majority of individuals whose data is shared.

154.   These challenges indicate that further discussion and engagement through the open policy making process is required to ensure that the intent of the term is appropriately captured.

Policy rationale

155.   The power would meet the objective of facilitating data sharing where it would directly benefit service recipients, by enabling agencies to better tailor services, but also protects privacy by restricting the agencies and the purposes involved in any particular share quite tightly. It is sufficiently future proofed to meet data sharing needs of public policy delivery in the future, while ensuring that the privacy of individuals remains paramount. Please note that this does not include data sharing for the purposes of punitive measures such as criminal prosecution.

156.   This approach fits well with the Data Protection Act and would support organisations to meet the fair processing and transparency requirements of the DPA. The ICO's Data Sharing Code of Practice emphasises the importance of having a clear purpose or objective (or set of objectives) for sharing data. Being clear about what the sharing is meant to achieve can also help organisations decide what data they need to share, with whom and whether the sharing is justified and proportionate. Data sharing arrangements can then be designed with this purpose in mind.

157.	This power is intended to be used in situations where:

- The objective could not be met without data sharing.
- It is *not realistic and practicable* to use consent to achieve the intended outcome or use of consent would not meet the criteria of free and informed decision.
- Analysis of anonymised data would *not achieve the intended* outcome.

For more detail on the process see figure 1.

158.	The policy is designed in such a way that it can be amended through secondary legislation in the future to reflect changes in social need and in social policy.

Other options identified and appraised

159.	A very broad power that would enable public agencies to share data when it is in "the public interest." However, this would not adequately balance needs of privacy and the protection of individual rights with the public interest and as a result there was no consensus on this approach.

160.	Specific gateways that provide the legal ability for public agencies to share information on a case by case basis. However, there are three key disadvantages to this:

- The time, cost and political will required to legislate means that government will always lag a few years behind meeting the data sharing need identified to improve service to citizens and in some cases the need will not be met at all;

- A case by case approach leaves the development of safeguards up to individual cases; this limits the ability of government to set up a robust framework of safeguards that will be applied consistently to data shares, as would happen with the recommended option; and

- It would not meet the key objective of future proofing the policy.

161.	A power which is flexible and enables data sharing between public bodies around particular population groups (instead of by objective) such as: households with multiple disadvantages in England; Ex-offenders; 16-19 year olds whose activity isn't known to the local authority and are at risk of becoming NEET. This approach does meet the criteria of flexibility balanced with constraint, but there are three key disadvantages:

- The risk of stigmatising or labelling people according to a service for which they are eligible;

- Difficulties of trying to define certain categories of people in law, particularly where there isn't an existing definition; and

- Difficulties if the policy aims to support people from more than one group.

162. No legal change to the current situation and a focus on cultural change. This would not address the legal barriers to sharing data to tailor public services better to individuals and therefore not solve the problems identified. This programme of work will supplement the cultural change work in local places lead by the Centre of Excellence for Information Sharing.

Safeguards

163. Even where a permissive power is conferred, a final decision by an agency to share data would be subject to DPA, art 8 ECHR, any other confidentiality requirements set out in EU law, any additional protocols/safeguards imposed across the CO's data sharing proposals and have regard to statutory codes of practice and the department's own guidelines.

164. The interaction between the new power and existing statutory gateways and frameworks (e.g. CRCA for HMRC) are to be considered further.

Method of delivery

The process for sharing TPS Data using the flexible gateway

165. Figure 1 demonstrates the process for deciding to use the new power, and how this relates to the other work streams in the broader data sharing work.

166. Department A seeks data held by Department B. The proposed data share meets the criteria of the TPS power and therefore there is a legal gateway. Department A sends a business case to department B.[14] It is the prerogative of Department B to agree, or not, to Department A's proposal.

167. Key elements for inclusion in the business case are:

- The entire route of the proposed data share, e.g. from Department B to Department A, further internal disclosure/matching within Department A, and any onward disclosure[15] to Local Authority C or Agency D;
- How it meets the legal gateway criteria i.e. <u>how the proposed data share would benefit the end user</u>;
- Costs and how these would be divided; and
- Fit with Government and Departmental priorities, i.e. *departments/authorities involved should avoid taking action which would be contrary to core business/functions*.
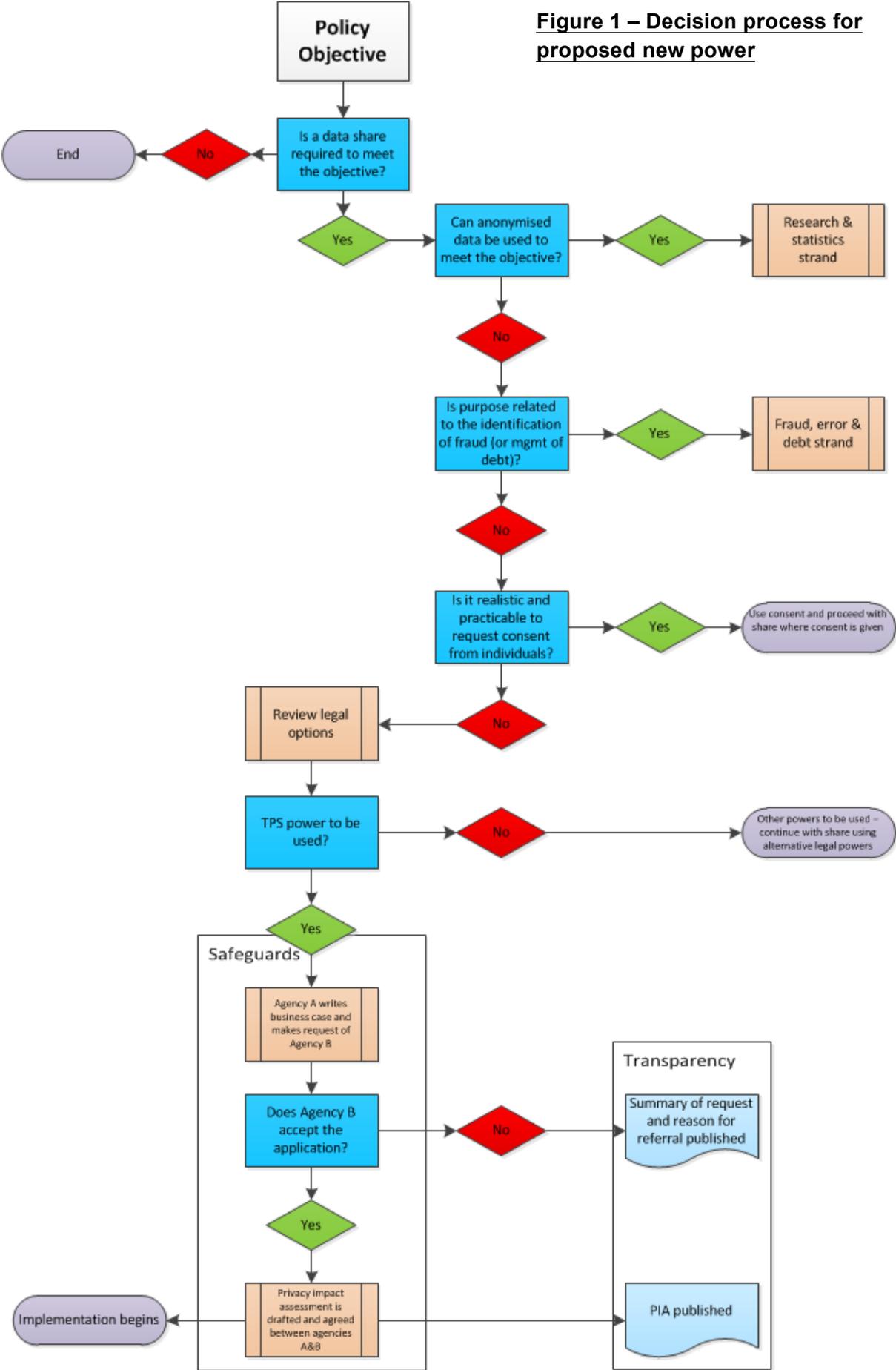
168.   The decision making processes within departments/agencies are up to the department /agency, but the decision as to whether to agree to a data sharing process using the TPS power should be taken at a level of seniority on or above the agreed minimum level. The recommendation is that a Senior Responsible Officer (SRO) for data sharing be appointed at Board level. Departments/public agencies currently have their own processes for managing information and making decisions about data sharing.

169.   Department B, as the data controller, would also make any decision to share data conditional on certain requirements being met. It is expected that these would include universal/internal standards for data and information storage, disposal, governance and assurance. It is recommended that data shares are subject to a statutory duty on all parties "to comply with up-to-date industry-recognised standards and best practice" or similar. All parties would be expected to comply with Principle 7 of the DPA.

170.   Once agreement has been reached in principle, parties will proceed towards drafting the Information Sharing Agreement. Key supporting documents will include Impact Assessments, the most relevant here being the Privacy Impact Assessment (PIA).

---

[15] Note on onward disclosure below

Figure 1 – Decision process for proposed new power

Policy Objective

Is a data share required to meet the objective? — No → End

Yes → Can anonymised data be used to meet the objective? — Yes → Research & statistics strand

No → Is purpose related to the identification of fraud (or mgmt of debt)? — Yes → Fraud, error & debt strand

No → Is it realistic and practicable to request consent from individuals? — Yes → Use consent and proceed with share where consent is given

No → Review legal options → TPS power to be used? — No → Other powers to be used – continue with share using alternative legal powers

Yes → Safeguards

Agency A writes business case and makes request of Agency B → Does Agency B accept the application? — No → Transparency: Summary of request and reason for referral published

Yes → Privacy impact assessment is drafted and agreed between agencies A&B → Implementation begins / PIA published

Thematic issues

Criminal Offences and other penalties

171.	Criminal offence for unlawful disclosure of personal information (similar to CRCA 2005, DWP legislation such as Social Security Administration Act 1992 s123, and SRSA 2007) would apply to all bodies when using the power. DPA section 55 would also apply to any people knowingly or recklessly obtaining or disclosing personal data, as would s.55A DPA (the IC's power to serve a monetary penalty notice on a data controller for a serious contravention of the data protection principles.

Oversight of the data sharing process

172.	Currently Parliamentary committees consider data sharing issues on a 'subject based' basis as they arise. This oversight can range from detailed inquiries into data handling such as the Health Committee's inquiries into the handling of NHS patient data to more specific recommendations on data sharing by Select Committees as part of wider reports into specific programmes, for example the Work and Pension Select Committee's report on Universal Credit and the Energy and Climate Change Committee's report on fuel poverty. The Public Accounts Committee also makes recommendations on data sharing on cross government initiatives for example in its report into programmes to help families facing multiple challenges.

Transparency

173.	It is proposed that the following be published two weeks ahead of the final decision being made on implementation:

- A list of data shares agreed under the power, and a summary of their purpose; and
- The PIA produced for each share – see ICO website for further information on PIAs.

Future Proofing

174.	To make the power flexible and future proofed, the policy objectives and the public bodies will be amendable through secondary legislation going forward.

175.	The proposal is that this will be amendable through secondary legislation that will need to be signed off by the Secretary of State. The recommendation is a minimum period of consultation and an affirmative process in Parliament.

176. Data sharing is intrinsically linked to changes in technological capability. For this reason it is recommended that a regular review of the impact of new technology on data sharing safeguards be instigated, and where necessary safeguards added/ amended in order to ensure that the policy continues to deliver the intended objective in light of technological advances.

Quality of data

177. The DPA already places a statutory duty on all data controllers to take adequate steps to ensure that data is accurate and where necessary up-to-date.

Onward disclosure /Use of data beyond the purpose for which it was shared

178. The power would only permit onward disclosure with the permission of the originating department; and only where a lawful power existed for that onward disclosure and it complied fully with all applicable DPA and HRA requirements.

Additional protections of personal data:

179. The following key measures are recommended to further protect personal data:

- Where relevant trials/pilots will be used to ensure the data sharing or linking meets identified objective. As noted above, this is a chance to test whether the criteria for the share can be met, as well as to explore if the policy objective can be achieved through the data match. If the pilot does not demonstrate the expected benefit then a broader data share will not be rolled out;

- Where practicable and realistic consent will be used and individuals will be asked permission to share their personal data with other agencies (e.g. older people and people with disabilities);

- Personal data will only be used where fully anonymised data would not meet the objective; and

- Minimisation of data shared. In every case departments will ensure that the personal data which are shared are limited to those which are necessary to achieve the objectives.

Next steps

- Further work to refine the wording of the policy, to ensure that it meets the twin goals of supporting improved service delivery and protecting personal data.
- Further work to identify case studies against which to test proposals.

Part 3 – Conclusions and Next Steps

180. The recommendations set out in this paper are a result of an open and productive dialogue between government officials and civil society and privacy organisations. Where the rationale for change has not been sufficiently robust, such as with initial proposals relating to fraud, strong challenge by civil society groups have shaped recommendations that aim to better understand the problem and the value of intervention. Where the case has been compelling, such as with tailored public services, it has been some of the active participants from civil society organisations from the OPM group that have pushed government to go further in defining how broad a power should be to provide longevity. These are strong indicators that the process is working and that all those participating are listening and shaping the work.

181. These findings represent the first stage of the process. In some instances more OPM work needs to be carried out on unresolved components of recommended proposals. These include:

   - Further work to refine the wording of the tailored public services policy option, to ensure that it meets the twin goals of supporting improved service delivery and protecting personal data.
   - Further work to identify case studies against which to test proposals.
   - Further discussion around the accreditation body under the de-identified data proposal;
   - Further consideration of alternative options to Affirmative Resolution safeguards for identified data within the research and statistics proposal; and
   - Further work to define debt and how it is understood and segmented across government.

182. Though methods of delivery have been set out where possible for each of the strands, more engagement through the OPM process is required to develop the details of implementation. It is essential that considerations around implementation are fully factored in the design of policy proposals. The OPM process will continue until the end of 2014, at which point the findings will be reviewed and next steps outlined as appropriate.

183. The final summary document will capture and summarise options identified by devolved administrations as part of their own OPM processes on the three strands. We will consider their findings and identify where there is scope for cooperation to achieve agreed outcomes.

184. An indicative outline timetable of key actions and dates is set out in Table 1.

Table 1 – Timetable of key actions		
Activity	**Date**	**Action owners**
Research and Statistics		
De-identified data strand		
• Development of proposals of what powers will be in primary and secondary legislation. • Health exclusion from power	6/10/14 to 7/11/14	Cabinet Office
• Proposals reviewed by OPM group (by correspondence)	23/10/14 to 7/11/14	OPM group
• Proposals revised in light of comments from OPM Group	10/11/14 to 14/11/14	Cabinet Office
Identified data strand		
• Developing further details on governance/approvals process for alternatives to Parliamentary scrutiny	06/10/14 to 5/11/14	ONS
• Consideration of governance/approvals process by OPM group	5/11/14 to 19/11/14	OPM Group
• Governance and approvals process revised and further developed in light of comments from OPM Group (by correspondence – but a meeting is possible if requested)	19/11/14 to 3/12/14	ONS and Cabinet Office
Fraud, Error and Debt		
• Further work to define debt and how it is understood and segmented across government with a view to determining whether it remains within scope of this work.	27/10/14 to 07/11/14	Cabinet Office (in collaboration with OPM Group)
Social media research		
• Social media research implementation	01/10/14 to 31/10/14	Cabinet Office
• Evaluation of social media research	03/11/14 to 07/11/14	Cabinet Office
External survey		
• Design external survey	22/10/14 to 24/10/14	Cabinet Office
• Review by OPM group (via correspondence)	27/10/14 to 31/10/14	OPM group
• Commission and implement survey	03/11/14 to 28/11/14	Cabinet Office
• Evaluate survey findings	15/12/14 to 19/12/14	Cabinet Office
Internal survey		
• Design internal survey	27/10/14 to 29/10/14	Cabinet Office
• Commission and implement survey	30/10/14 to 28/11/14	Cabinet Office
• Evaluate survey findings	01/12/14 to	Cabinet Office

		05/12/14	
Studies			
•	Commission case studies from OGDs and CSOs on identified gaps	w/c 03/11/14	Cabinet Office
•	Case studies provided by OGDs and CSOs	10/11/14 to 05/12/14	OPM Group
•	Evaluation of information from case studies	w/c 08/12/14	Cabinet Office
Pilots			
Error pilot			
•	Design Error pilots	20/10/14 to 31/10/14	Cabinet Office
•	Review of design by OPM Group	03/11/14 to 07/11/14	OPM Group
•	Implement error pilot	10/11/14 to 05/12/14	Cabinet Office
•	Evaluate pilot	08/12/14 to 12/12/14	Cabinet Office
Risk Based Approach Pilot			
Dummy data			
•	First design of methodology	13/10/14 to 31/10/14	Cabinet Office
•	OPM review of methodology	03/11/14 to 07/11/14	OPM Group
•	Implement Pilot with dummy data	10/11/14 to 05/12/14	Cabinet Office
•	Assessment of Pilot	08/12/14 to 12/12/14	Cabinet Office
Live Data			
•	First design of methodology	24/11/14 to 5/12/14	Cabinet Office
•	OPM review of methodology	08/12/14 to 12/12/14	OPM Group
•	Implement pilot with live data	15/12/14 to 23/01/15	Cabinet Office
•	Assessment of pilot	26/01/15 to 30/01/15	Cabinet Office
Tailored Public Services			
•	Further work to define powers	13/10/14 to 14/11/14	Cabinet Office
•	OPM session on detailed wording of proposed power and objectives	w/c 17/11/14	OPM Group
•	Working group OPM session on detailed wording	w/c 17/11/14	OPM Group
•	Policy option redrafting following OPM group comments	24/11/14 to 28/11/14	Cabinet Office
•	OPM review of redrafted paper	01/12/14 to 05/12/14	OPM Group
Overarching activities			
•	Phase 1 Plenary Session	22/10/14	OPM Group
•	Mid-Phase 2 Plenary Session	02/12/14	OPM Group
•	Final paper circulated to OPM Group for review	w/c 08/12/14	OPM Group

	Early January 2015	
• Final Plenary Session	Early January 2015	
• Release of paper	January 2015	Cabinet Office