Research Findings
No. 19/2012

social research

Public Services and Government

Public Acceptability of Cross-sectoral Data Linkage

Dr Sara Davidson and Christopher McLean, Ipsos MORI Scotland, with Professor Sarah Cunningham-Burley and Dr Claudia Pagliari, Centre for Population Health Sciences, University of Edinburgh.

To supplement the written consultation, A Scotland-wide Data Linkage Framework for Statistics and Research, the Scottish Government commissioned Ipsos MORI, along with Professor Sarah Cunningham-Burley and Dr Claudia Pagliari from the Centre for Population Health Sciences at the University of Edinburgh, to undertake a series of public deliberative events. The overall aim of the events was to explore the acceptability of linking personal data for statistical and research purposes, thereby identifying particular sensitivities and exploring mechanisms for overcoming concerns.

Main Findings

- All participants recognised potential benefits of data linkage but most also had questions and concerns about it. There was a view that linkage could lead to increased negative "labelling" of people, and thus discriminatory treatment and stigma, due to the potential for labels to carry across sectoral boundaries. There was also concern about the possibility of commercial organisations gaining access to linked data.
- A number of participants contended that linkage would increase the likelihood of data security breaches because more people would have access to more data and hackers would be able to obtain a significant amount of information about individuals "in one hit". Often these concerns were based on a mistaken assumption that linkage would result in the creation of one super-database of information that would be 'warehoused'.
- Participants often assumed that linked data would include personal identifiers. When they were told that this would generally not be the case, many became more comfortable with, or positively disposed towards, linkage. However, a minority contended that anonymised data could always be linked back to personal identifiers, by anyone with the necessary knowhow.
- Participants broadly supported the objectives of the Data Linkage Framework but sought reassurance on: who would oversee its operation; who would have access to linked data; how privacy would be protected and data kept secure; and where accountability would lie if data was lost or stolen. The draft Guiding Principles were seen to go some way towards providing this reassurance but they were also often considered too vague and open to interpretation by vested interests.
- Suggested safeguards to maximise public confidence in the Framework included: publishing all processes and procedures surrounding data linkage, and details of who is using linked data; establishing an oversight body, comprising highly qualified professionals and, potentially, lay members, with responsibility for ensuring that the Principles are upheld; ensuring that explicit consent is obtained for uses of data containing direct identifiers and setting clear parameters around what is being consented to; ensuring all electronic systems used by individuals/organisations with access to linked data meet a minimum security requirement that is reviewed and updated frequently; ensuring all researchers and officials with access to linked data are appropriately vetted; imposing strict sanctions on individuals/organisations responsible for any breaches of the Principles and specifying the range of possible sanctions within the Principles.

The Scottish Government

Research aim and objectives

The overall aim of the events was to explore the views of the public on the acceptability of linking personal data for statistical and research purposes, thereby identifying particular sensitivities and potential barriers to public confidence and exploring mechanisms for overcoming concerns.

Specific objectives were to:

- identify particular concerns or sensitivities around the sharing and linking of data within and between sectors
- identify whether any particular sector-to-sector linkages raise levels of concern about privacy
- test the extent to which the draft 'Guiding Principles' reassure participants that data linkage will be governed appropriately
- identify what safeguards could be put in place to maximise public confidence
- investigate the extent to which the public support the objectives of the 'Beyond 2011'[1] project and the extent to which these raise further privacy concerns
- investigate whether the public have views about ongoing public involvement and how this might be achieved.

Methodology

The study was conducted using deliberative methods in recognition of the complexity and potential unfamiliarity of the topic, and thus the need for participants to be appropriately informed in order to meaningfully consider the issues.

Three half day workshops were held; in Stirling, Inverness and Glasgow, between 26 May and 9 June 2012, with participants recruited to be broadly representative of the Scottish population.

Thirty participants were recruited for each workshop, with the aim of ensuring that around 25 attended. In the event, 24 attended the Stirling event, 22 attended the Inverness event and 27 the Glasgow event. Attendees were representative of the wider pool of recruits.

[1] A project being run by the National Records of Scotland to assess alternative options for producing population and socio-demographic statistics, including the use of administrative data sources.

Findings

General attitudes to organisations holding and using data about individuals

When participants were asked how they felt about organisations holding information about them, they tended to begin by expressing concerns about an encroaching "Big Brother" or "surveillance society" and/or about the amount of data on individuals that is collected and used in the commercial sphere.

The term "big brother" society was used to refer generally to the large amount of data that is collected on individuals (across both the public and commercial spheres), and also to a proliferation of surveillance mechanisms such as CCTV, electronic tagging and mobile phone tracking. The overall feeling was that the monitoring and recording of different aspects of people's lives has gone "a bit far", compromising privacy.

Most participants were acutely aware and strongly critical of the tendency for commercial actors to sell individuals' details to each other for use in targeted marketing campaigns. Several also expressed concern about the growing ease with which such actors can "profile" individuals by drawing on data from multiple and diverse sources, including social networking sites like Facebook. There was a perception that this provides fertile ground for scams, fraud and identity theft.

Spontaneous detailed comments about the holding of personal information by public bodies were sometimes slower to emerge – or at least less well formed – than those relating to commercial actors. Nevertheless, when prompted for their views, virtually all participants engaged keenly with the subject.

Many said that they generally trusted public bodies more than commercial organisations with their personal data. The NHS tended to be seen as particularly trustworthy due to the fact that health professionals are expected to abide by a moral code of conduct as part of their job and, more generally, to serve or help the public.

Still, among a significant minority of participants, there was scepticism around the extent to which public bodies could be trusted to look after data and use it appropriately. This was in part fuelled by high profile cases of data losses and data breaches, and a perception that public bodies are active in selling data to commercial organisations.

Many participants were also conscious of the fallibility of information technology and the difficulty of creating entirely "fool proof" security systems and

procedures. There was particular concern about the potential for hacking, which tended to reflect personal experiences of fraud and other scams in the commercial domain, as well as media stories of security breaches in large organisations, such as LinkedIn.

Other participants spoke less about the fallibility of systems and more about security risks arising from the "human factor" in public bodies; specifically, the potential for "human error" in data handling and for officials to behave indiscreetly, unscrupulously or corruptly.

While most participants acknowledged that public bodies need to hold data on individuals, some contended that they do not always seem to make effective use of the information they have. Paradoxically, others felt that there is too much focus on (quantitative) research data and statistics in decision making, with the effect that individuals and groups are often crudely categorised and "labelled" – a practice that was seen to result in discriminatory treatment and/or stigma, as well as policies and spending plans that fail to reflect the myriad of ways in with social and other problems are experienced across the population.

Unprompted views around cross-sectoral data linkage

Unprompted views around cross-sectoral data linkage were sought prior to an informational presentation about the proposed Data Linkage Framework.

While virtually all participants recognised potential benefits of data linkage, most also had questions and concerns about it. A common concern was that it could lead to increased negative "labelling" of people. More specifically, there was concern about the potential for labels to carry across sectoral boundaries and result in individuals or groups experiencing discriminatory treatment or stigma in multiple spheres.

Concern was similarly expressed over the possibility of commercial or political actors gaining access to linked data. There was a strong consensus that this would not be acceptable.

A number of participants contended that linkage would increase the likelihood of data security breaches, both because more people would have access to more data, and hackers would be able to obtain a significant amount of information about individuals "in one hit". Often these concerns were based on a mistaken assumption that linkage would result in the creation of one super-database of information that would be 'warehoused' for use by multiple organisations.

Participants often assumed that linked data would include personal identifiers. When they were told that this would generally not be the case, many became more comfortable with, or positively disposed towards, linkage. However, a small number of participants were keen to emphasise that, even with anonymisation, there would be the potential for groups to be negatively profiled and labelled. Other, particularly IT literate, participants contended that anonymised data could always be linked back to personal identifiers by anyone with the necessary knowhow.

The Data Linkage Framework

Participants broadly supported the overarching objectives of the Data Linkage Framework. However, they had specific concerns on which they sought reassurance, which tended to centre around the questions of:

- who would oversee the operation of the Framework
- who would have access to linked data, and specifically whether this would include commercial companies
- how individuals' privacy would be protected
- how the data would be kept secure
- where overall accountability would lie if linked data was lost or stolen

These concerns were discussed in relation to data linkage generally and there was little explicit differentiation between particular sector-to-sector linkages, despite prompting by the facilitators.

Perceptions of the draft Guiding Principles were mixed. On one hand, there was a view that the Principles go some way to addressing the main areas of concern and provide reassurance that data linkage will be carried out appropriately and securely. On the other hand, the Principles were commonly considered to be too "vague" and therefore open to interpretation and manipulation by vested interests. In particular, it was felt that there was a lack of detail surrounding who would be on the oversight body and the sanctions that would be imposed for breaching the Principles.

Participants identified a number of safeguards that could be implemented to maximise public confidence in the Framework. These included:

- a requirement that anyone applying to use linked data must provide a strong justification to a commission or panel as to why their research is in the public interest
- publishing all processes and procedures surrounding data linkage, as well as details of who is undertaking research using linked data

- establishing an oversight body, comprising highly qualified professionals and, potentially, lay members, with responsibility for granting or refusing data linkage requests, ensuring that the Principles are upheld, and administering sanctions
- establishing accountability by placing data linkage under ministerial remit or the auspices of an independent professional or senior civil servant
- ensuring that explicit consent is obtained for uses of data containing direct identifiers and that, in the process, clear parameters are set around what is being consented to
- requiring that this consent be obtained from data subjects themselves or from their next of kin, and preventing any oversight body from granting proxy consent
- ensuring all electronic systems used by individuals and organisations with access to linked data meet a minimum security requirement that is reviewed and updated frequently
- ensuring all researchers and officials with access to linked data are appropriately vetted through mechanisms such as: a certified training course; an accreditation scheme; or an assessment scheme similar to Disclosure Scotland.
- imposing strict sanctions on individuals and/or organisations responsible for any breaches of the Principles and specifying the range of possible sanctions within the Principles

There was broad support for the objectives of *Beyond 2011*, which did not raise any new privacy issues. Generally, 10 years was considered too long a gap between censuses and several participants questioned whether the census represents value for money given that the data soon becomes obsolete.

There was a strong appetite for ongoing public engagement in the development of the Data Linkage Framework and in particular for media advertising campaigns; the distribution of informational leaflets; and establishment of a dedicated website that could serve as a 'one stop shop' for everything members of the public might want to know about data linkage.

This document, along with full research report of the project, and further information about social and policy research commissioned and published on behalf of the Scottish Government, can be viewed on the Internet at: http://www.scotland.gov.uk/socialresearch. If you have any further queries about social research, please contact us at socialresearch@scotland.gsi.gov.uk or on 0131-244 7560.

Further information on the Data Linkage Framework can be found at: http://www.scotland.gov.uk/Topics/Statistics/datalinkageframework

ISBN: 978-1-78256-013-5

APS Group Scotland
DPPAS13303 (08/12)

GSR
GOVERNMENT SOCIAL RESEARCH
Social Science in Government