

Can data-sharing improve public services? Lessons for Parliament

 blogs.lse.ac.uk/politicsandpolicy/can-data-sharing-improve-public-services-some-lessons-for-parliament/

11/24/2016

The Digital Economy Bill, currently passing through Parliament, includes proposals for HMRC information on benefits recipients to be shared with the Department of Energy and Climate Change, in order to identify citizens living in fuel poverty. Sharing data between government departments for policy purposes is not as straightforward, explains [Edgar Whitley](#), and outlines some of the key issues that must be kept in mind.



It is easy to see the appeal of data-sharing as means of fixing problems with the delivery of public services. Take data set X held in one part of government, share it with a different part of government that has another data set Y. Combine data sets X and Y and use the results to achieve a particular policy objective.

A current example of data-sharing, found in Part 5 of the [Digital Economy Bill](#) currently being debated in Parliament, aims to address fuel poverty. The idea is to identify those citizens living in fuel poverty by sharing tax credit (benefits) data held by HMRC and combining it with basic property characteristics data held by the valuation office agency and the Department of Energy and Climate Change (DECC). Having identified those citizens who would benefit from targeted assistance, the final stage is to inform licensed energy providers about which of their customers should automatically receive the assistance.

This form of data-sharing depends on matching records in data set X with records in data set Y and requires high quality data to avoid type 1 (spurious match) and type 2 (undetected match) errors. Ensuring high quality data is a non-trivial managerial task often involving misaligned incentives. If managed badly, the adage “garbage in, garbage out” applies and fixing this can add considerable “hidden” costs to any data-sharing activity. In addition, there is a clear need for legal and technical safeguards as the sensitive personal data of citizens.

In an [article](#) published nearly 30 years ago, Australian academic Roger Clarke highlighted the problems that can arise when data-sharing and matching is based on poor quality data. He reported that when the US Department of Health, Education and Welfare ran its welfare files against its own payroll files, 33,000 potential “matches” were found. A year’s worth of further investigation reduced this to a narrower set of 638 cases. Of these only 55 were prosecuted. Moreover, when 15 of these cases were independently reviewed, 5 were dismissed, 6 were convicted of felonies and no prison sentences resulted (pp. 508-509).

In 2016, a review of another data-sharing agreement revealed many similar trends. The [report](#) by the Independent Chief Inspector of Borders and Immigration examined the implementation of a data-sharing agreement that was intended to help create a “hostile environment” for individuals who are in the UK without valid leave.

Two specific measures were examined: the refusal by the Driver and Vehicle Licensing Agency (DVLA) of applications for a UK driving licence and the revocation of existing licences for individuals not lawfully resident in the UK; and, the requirement placed on banks and building societies to refuse an application for a UK current account from an individual listed as a 'disqualified person'. Both processes rely on bulk data-sharing between the Home Office, DVLA, and Cifas were based on pre-existing collaborative arrangements between the different organisations and, of course, rely on the quality of the (Home Office) data being shared.



The inspection found, however, that “records for individuals were incomplete or had been completed incorrectly (with data placed in the wrong fields), or there were delays in updating records” (2.7). To mitigate these problems “the Home Office checks its records manually for new licence applicants and also to confirm that the DVLA should proceed with a revocation” (2.8).

Even in the cases where individuals were correctly identified and had their driving licence revoked, only a small proportion surrendered their revoked licence. This undermines the intended policy objectives: “stopping illegal migrants from being able to drive lawfully, and from using a driving licence to access other benefits and services”. This latter point might have particular significance if it becomes necessary to present government issued photo ID documents (passport/driving licence) before [accessing health services](#).

Similar data quality issues affected the bank account opening checks, with 10 per cent of a sample of 169 reviewed cases incorrectly listing people as “disqualified persons” [i.e. who shouldn't be allowed to open a bank account] but who in fact either had leave to remain in the UK or outstanding applications / appeals (6.29).

The report found problems with the safeguards that had been put in place including the memoranda of understanding about data-sharing between the various agencies (2.6). Moreover, the inspector's concerns that the justification for the policy “is based on the conviction that they are ‘right’ in principle, and enjoy broad public support, rather than on any evidence that the measures already introduced are working or needed to be strengthened” (2.21) has direct parallels with Clarke's concluding point that the benefits of data-sharing “must not be assumed, but carefully assessed” (p. 511).

As Parliamentarians discuss the data-sharing proposals, let us hope that they heed these lessons.

Note: background Parliamentary submissions by the author which the above draws on can be found [here](#), [here](#) and [here](#).

About the Author

Edgar Whitley is Associate Professor of Information Systems and Director of Teaching and Learning at the Department of Management, LSE.

