



Llywodraeth Cymru
Welsh Government

www.cymru.gov.uk

Recommendations for Improving Research Access to Potentially Disclosive Data

Research Summary

Social research

Number: 26/2014

Aim and Scope

1. A fellowship on Improving Research Access to Potentially Disclosive Data for Wales was funded jointly by the Economic and Social Research Council (ESRC) and the Welsh Government (WG) as one of four fellowships created to inform the WG's Programme to Maximise the Use of Existing Data.
2. The aim of the fellowship was to:
 - identify stakeholder needs with regard to access to detailed administrative and survey data for research purposes;
 - review the existing options for making data available to researchers; and
 - make recommendations to WG on the best way to establish a technical and procedural infrastructure to maximise access to data on Wales.

3. The fellowship addressed data access for researchers both within and outside WG, taking into account situations where data must remain within WG for legal or practical reasons as well as those where data can be made available by specialist organisations outside WG.

Methodology

4. A stakeholder consultation was undertaken, consisting of face to face interviews and email and web feedback requests. Twenty-eight interviews were carried out with WG analysts and policymakers, academic researchers, data sharing specialists from other government departments, third sector researchers and managers of data access infrastructures. Feedback was sought from the Welsh Statistical Liaison Committee, from across WG via the intranet, from Welsh Universities and through the research fellow's "rlab-data" blog.

Key Findings

5. The project found that there was a definite demand for more data on Wales, particularly at a detailed geographical level. However, few of the stakeholders interviewed felt confident that they had the knowledge and the infrastructure to obtain access to the data they wanted. This problem is exacerbated by the complexity of UK data sharing legislation and the dispersion of expertise within WG Knowledge and Analytical Services (KAS). In many cases the data required are not owned by WG, so the issues are not only of access but also availability and therefore the challenge is not only in developing infrastructure but also data sharing procedures and legislation.
 - Much of the statistical and administrative data on Wales is collected or held by public bodies other than WG, in particular UK government departments.
 - UK legislation governing data sharing is very complex, making it difficult to identify appropriate legal mechanisms to enable WG to secure the data resources it needs from other public bodies.
 - Analysts in KAS can find it challenging to identify suitable data to support their work. Knowledge of the data resources available, access restrictions and legislation is spread throughout KAS and information can therefore be difficult to locate.
6. Key issues identified included:
 - A demand for more detailed data on Wales to support the evidence base across all stakeholders, including researchers from WG, academia and the third sector.
 - There is a need for data at a lower geographical level than 'Wales', particularly to support the evaluation of geographically targeted policy interventions.
 - Researchers both within government and outside reported that they did not have access to the appropriate secure infrastructure to support their data access needs.
7. Traditionally, access to data has been viewed as a straightforward trade-off between risk and utility, meaning that the more useful the data the greater the risk of releasing it. In recent years the data access infrastructure has changed significantly. Technical and procedural advances mean that even highly detailed data can be safeguarded to a level where the risk is significantly lower than previously. In 2006 a five-part security model 'the Five Safes' was designed to address data release at all levels of sensitivity and includes five components

that impact data security – safe projects, safe people, safe settings, safe data, safe outputs. Aspects of these five components can be traded off to ensure data security. This new framework offers the opportunity to provide access to sensitive data while maintaining data security in a way that supports the legal and ethical requirements to safeguard data. The framework was developed for the ONS and underpins much of the UK's secure data access infrastructure including the ONS VML, the UKDS secure service and the HMRC Datalab as well as drawing increasing interest internationally.

8. The UK has a long-standing tradition of microdata analysis to support policy based research, and government departments have been making data available in partnership with academia for over 40 years via the UK Data Service (and its predecessors). A major new ESRC investment in a 'Big Data Network' will take place in 2013 leading to a significant increase in opportunities available to UK researchers for processing previously underused microdata resources from administrative systems, social networking, local government and the third sector. This major investment in UK data access infrastructure offers WG the ideal opportunity to maximise access to data

resources and opportunities for research on Wales.

9. The most important component of any plan to maximise access to potentially disclosive data for Wales is a Data Management Policy (DMP). A formal DMP would provide the foundation on which all other components of a data access strategy could be built. It would offer a comprehensive overview of holdings and a framework for documentation and discovery it would also increase data security by formalising data classification, storage, preservation and sharing.

Key recommendations:

- Implement a Data Management Policy to enable all WG data resources to be recorded, archived and documented and where possible
- Establish a KAS Data Service to increase the efficiency of data users within WG.
- Develop a network of experts to provide bespoke analyses of potentially disclosive data.
- Require all contracts with external organisations undertaking data collection for WG to include clauses specifying that microdata must be deposited with WG to facilitate quality control, reuse and where possible linking.
- Make full use of existing infrastructure, and developing standards nationally and

internationally in order to minimise expenditure, leverage existing user support, promote standards, maximise access to resources and streamline access.

- Implement a Safe Setting to maximise access to external Research Data Centres and the potential for secure in-house processing of data held by WG.
- Investigate the costs and benefits of developing a secure data access infrastructure in-house.
- Continue to monitor and engage with national and international projects to maximise access to funding opportunities and to increase the chances of forthcoming data projects meeting WG needs.

Implementation

10. In the current financial climate it may be a challenge to find funding to implement the above recommendations. Any efforts to increase access to data resources will necessarily attract a cost, primarily in staff time to implement changes and to develop and maintain systems. Additional investment in technology and, for safe settings, space will be required. In the medium to long term the cost of maintaining systems and providing data support also has to be considered.

11. The key to minimising costs is to make use of existing services and infrastructures as far as possible, as recommended above. Where suitable resources do not exist WG should consider partnering with funding councils, institutions, other governmental organisations charities and, where appropriate the commercial sector to develop systems, services and datasets of mutual benefit. Despite the unavoidable costs the potential for greater efficiency and saving for WG in the medium to long term is significant. For example, effective implementation of the recommendations might be expected to reduce the need for future spending on primary data collection, data processing and analysis; reduce the burden on survey respondents thus saving time and money for the public and businesses in Wales; and improve policy impact due to an increase in the range and quality of the evidence base.

12. In terms of efficiency a centralised KAS data management service overseeing a formal Data Management Policy could contribute to:

- An increase in expertise available to KAS (and WG more widely) through the investment in staff specialising in issues relating to data use, analysis, and acquisition.

- An improvement in the quality of research by providing expert support for the analysis of data on Wales.
- A reduction in the lead in time to research by assisting with negotiations for access to external resources; and by facilitating data discovery and implementing transparent access procedures for internal resources
- A reduction in the time needed to access data-related information and resources by facilitating WG's ability to locate and consult expert staff.

13. Adoption of the recommendations in this paper would enable KAS to significantly advance the range and quality of resources available to researchers interested in carrying out evidence based research on Wales.

Views expressed in this report are those of the researchers and not necessarily those of the Welsh Government

Author:

Tanvi Desai,
London School of Economics
Research Fellow, Welsh Government

Contact:

Sarah Lowe
Phone: 029 2082 6229
Email: sarah.lowe@wales.gsi.gov.uk

ISBN: 978-1-4734-1085-5

© Crown Copyright 2014