

Towards Government 2.0

DATA as a SERVICE

INTRODUCTION

Much of the Government 2.0 movement is defined around two key facets – the delivery of more Public Sector Information (Data Engagement) and the delivery of improved online communication both with and across government agencies (Social Engagement). Individually, these two facets can each help with delivering consultative, informed decision making between citizens and their government. However, real collaborative Government 2.0 engagement is most likely to be achieved where Data Engagement and Social Engagement are successfully balanced and integrated together.

Data Engagement must provide timely, open, trusted and confidentialised data. Information at all levels, based on micro data and its relevant metadata, should be easily accessible on public web sites using intuitive ad-hoc data dissemination and visualisation tools.

Communities of interest that form around accessible Public Sector Information (PSI) can generate effective Social Engagement with government. However, risk adverse public sector agencies often raise objections about releasing more public data, citing privacy protection as a constraint. Information dissemination technologies that embrace statistical disclosure controls can play an important role in removing these objections - providing more open access to data and building public trust and confidence in that data.

GOVERNMENT ENGAGEMENT - CULTURE + DATA

The charge of the Taskforce is to find ways to accelerate the development of Government 2.0 Engagement — to consult, collaborate, and open up government to public access. The Taskforce's goals can be paraphrased as promoting:

- Accessible, useable Public Sector Information (that is, data)
- Consultative, participatory government (enabled via online engagement)
- A culture of online innovation
- Collaboration across agencies.

These goals can broadly be grouped into the two contexts of promoting a collaborative, participatory culture and delivering access to the Public Sector Information that feeds Government policy and decision making.

Culture, Collaboration, and Participatory Government

The Culture and Collaboration goals are challenges for government leadership and innovation. Solutions to the challenges can be aided by the use of technology and data standards. Consultative, participatory government is clearly going to be supported by Web 2.0 software and services. Wikis, blogs, social networks, video sharing and mashups are becoming tools for public sector **Social Engagement**.

Accessible, Useable Government Data

Data Engagement is about the goal of promoting and delivering accessible, useable Public Sector Information. Web 2.0 concepts and technology support an audience increasingly comfortable with self-service, high performance and graphically interesting data dissemination tools. However, accessible, useable data must also be safe with individuals' privacy protected.

Here in Australia the Australian Bureau of Statistics is at the forefront of safely delivering valuable public data through its innovative CDATA Online and, soon to be released, TableBuilder services¹. These services allow users to dynamically construct ad-hoc tabulations, graphs and maps from the 2006 Census and export the resultant answer sets for reuse elsewhere.

What makes CDATA Online and TableBuilder truly innovative is that, for the first time, we have public systems where the source data being accessed is held at the most granular level possible – down to the individual unit record or micro data level². The individual unit records are not returned to the user, but aggregations and tabulations based on those records are.

This innovation has been made possible by using a tabulation engine that introduces small random adjustments into the aggregated answer set delivered back to the user's browser. These adjustments (known as perturbation) do not affect the general statistical validity of

¹ See <http://www.abs.gov.au/CDATAOnline> and <http://www.abs.gov.au/tablebuilder>.

² Allowing the public to generate tabulations based on Census micro data was not provided in the past because of privacy concerns. Census data does not carry information that allows individuals to be directly identified. However, without proper disclosure controls in place, a person with access to micro data could theoretically identify or target individuals and households using careful cross tabulation of characteristics unique to those people.

large tabulations (which is how Census data is mostly used) but still protects the privacy of Census respondents.

These same concepts and techniques could help remove privacy objections to the dissemination of other large Public Sector data sets, particularly in the social welfare, health and education spheres. Not only might it enable new data sets to be exposed to the public, it could also increase the value of information derivable from already publicly available data sets. By directly basing ad-hoc enquiries on data derived from the micro data level you have a greater range of field combinations available for cross-tabulation compared to the pre-aggregated data often published today. This flexibility promotes new ways of looking at data, delivers new insights and creates more value from Public Sector Information.

But it isn't just the raw data where you find the value in Public Sector Information. It is also in the meaning (the **metadata**) that public sector agencies spend considerable effort and expense wrapping around their data. Metadata helps data prosumers (producer-consumers) and end users make more sophisticated decisions that lead to more valuable outcomes. It increases their understanding and their trust in the data, while also providing a "duty of care" that helps minimise the chances of people unwittingly making incorrect interpretations or use of the data.

The ABS TableBuilder service is a case in point for delivery of sophisticated metadata. Users of that application can combine data from different Census data tables in different ways. However, not all combinations are as valid as others. So the application has been designed to alert the user when specific field value combinations occur in a table by dynamically generating specific annotations as table footnotes.

It is important when disseminating Public Sector Information to find ways to maximise the "stickiness" of the metadata within which the raw data was originally wrapped. Stickiness will come from making it efficient and easy for systems to process the data and metadata together – which is where standards like SDMX and DDI can help³. Stickiness can also come from delivering the data and the metadata together in a socially engaging, visually appealing way. This can be achieved through such things as RESTful web technology that allows tabulations, graphs and maps incorporating metadata to be easily embedded into 3rd party social sites. Or, by the Data Service Provider delivering an integrated, collaborative

³ See <http://sdmx.org/> and http://www.opendatafoundation.org/papers/DDI_and_SDMX.pdf for more information on SDMX and DDI.

experience that directly promotes community-based decision making and participation, i.e., combining Data Engagement with Social Engagement. Whichever approach is used it needs to be a data engagement experience based on the twin values of disseminating data with *safety* and with *meaning*.

GOVERNMENT AS A DATA SERVICE PROVIDER

Data Engagement is successfully executed by public or commercial organisations that see themselves as Data Service Providers, delivering data and service that promote informed decision making. The ABS is a prime example of an organisation that sees its role clearly as a Data *Service* Provider. By its charter, the ABS is defined as a Data Provider and as an organisation its culture embraces Data Engagement (what National Statistics Offices call Dissemination).

Similar to the OECD Principles for disseminating public sector information, in our dealings with Data Service Providers we have found the following six considerations to be a useful framework for discussions of successful Data Engagement.

Data as a Service

Governments as Data Service Providers have the challenges of openness and confidentiality to fulfill their obligations to provide data in both raw and visualised forms.

The goals of a government Data Service Provider can be to fulfill a stated mission, generate revenue, provide their constituents with useful information, or ensure a return on investment on a valuable public data asset.

Trust in a Brand

Government organisations have a valuable brand to protect. The public's basic trust in a government agency is often paramount. Protecting a brand includes such things as ensuring appropriate privacy protection, copyright and open data agreements.

For Data Service Providers a brand can be enhanced by the user experience. An intuitive, no-training-necessary experience is imperative. A good data experience requires robust metadata and compelling visualisation.

Data Integrity

Data integrity means confidentiality and privacy protection, data quality, accuracy and completeness, including relevant data and metadata. Use of standards such as SDMX and DDI will be important components of data integrity programs.

Statistical disclosure control and duty of care technologies can play an important role in removing the objections that Government may

have in providing more open access to data and providing trust and confidence in the data.

Responsiveness

There should be no reason to make a trade-off between timeliness and data quality. Public data must be right and must be on time to be truly relevant.

Community

Sharing government data is about providing communities of users and the public with data for informed decision making and government participation.

Collaboration with and within these communities will be the essence of Government 2.0, resulting in effective public/private partnerships.

Efficiency and Cost

Providing better, more complete data to the public can be either cost effective or a cost burden. Organisations can substantially reduce costs in this area by making more self-service data available.

Using data standards such as SDMX and DDI can reduce data integration costs. The availability of modern Web 2.0 software solutions for dissemination and collaboration reduces initial implementation costs as well as longer term support costs.

CONCLUSION

Here in Australia technology is already in use that demonstrates you can provide accessible, safe dissemination of Public Sector Information. The Australian Bureau of Statistics' CData Online and TableBuilder solutions provide privacy protected access to information derived directly from micro data. By embedding disclosure control technologies into its services the ABS has ensured a Return on (Data) Investment to the Australian public for their substantial census and survey collections. More data with greater quality is available to more user communities with less manual support – without compromising the safety of the data being disseminated.

Integrating and balancing data engagement and social engagement solutions will provide a foundation to achieve the Task Force's goals of consultative, participatory government, collaboration across agencies and a culture of online innovation. Meeting the cultural and technical challenges of trust, data integrity, responsiveness and costs will enable government/citizen engagement to expand.

Australian government agencies, citizen groups and commercial solution providers can take up these challenges and together build a position of global leadership in deploying successful Government 2.0 solutions.

ABOUT SPACE-TIME RESEARCH

Space-Time Research (<http://www.spacetime.com>) is the developer of SuperSTAR software, a web-based software suite delivering analytics and visualisation with confidentialisation and disclosure controls. The ABS is using SuperSTAR as the basis for their CData Online and TableBuilder services.