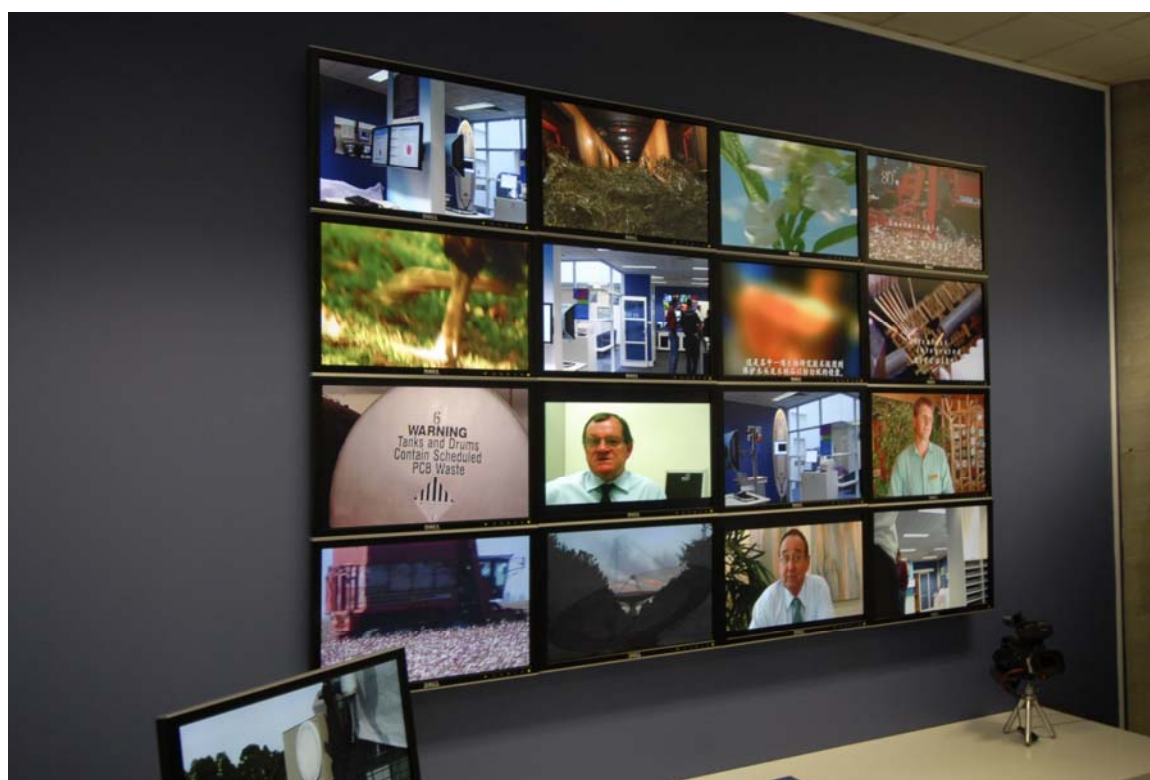


CSIRO Submission 08/313

Improving Access to Victorian Government Public Sector Information and Data

September 2008



Enquiries should be addressed to:

Mr Peter King
Senior Adviser
CSIRO Government Relations
PO BOX 225
Dickson ACT 2602
Australia
Email: PeterG.King@csiro.au
Ph: (02) 6276 6361

Main submission Author:

Dr Alex Zelinsky
Group Executive, CSIRO Information &
Communications S&T
PO Box 76
EPPING NSW 1710
Australia
Email: Alex.Zelinsky@csiro.au
Ph: (02) 9372 4202

Contents

Contents	2
Introduction	3
Responses to specific Terms of Reference	3
Terms of Reference A	3
Terms of Reference B	4
1. Open access to geo-spatial data:	4
2. Open access to Victorian health data:	4
Terms of Reference C	4
Terms of Reference D	5

Introduction

CSIRO welcomes the opportunity to respond to the Committee's inquiry. CSIRO considers the Inquiry is timely and is addressing an issue of increasing importance to all tiers of government and Australia's research sector. The response has particular regard to those aspects of the Committee's Inquiry which are particularly relevant to Australia's major research organisation and includes examples which demonstrate how increased access to government information would have clear benefits for national research outcomes.

More generally, CSIRO's response can be considered against the recently-released report of the Review of the National Innovation System, "*Venturous Australia – Building strength in innovation*", which makes a number of recommendations regarding open access to public information. Specifically, the report recommends that:

*"Australian governments should adopt international standards of open publishing as far as possible. Material released for public information by Australian governments should be released under a creative commons licence (Recommendation 7.8)"*¹

In general CSIRO supports the proposal for greater public access to information created by government-funded research, with this always needing to be subject to necessary conditions to ensure scientific independence and integrity. CSIRO is willing to work with Australian governments to effectively adopt standards for open publishing.

In making this response, CSIRO observes that the end-users of government-funded research have different requirements (in relation to public access or restrictions on access to information) and that these need to be taken into consideration. As a mission-directed R&D organisation, CSIRO has a requirement to keep a clear focus on the requirements to achieve impact from its research. In its partnerships with both the public and private sectors, CSIRO will need to be cognisant of the requirements of its partners and that are necessary to achieve the intended impact from the research, and on that basis determine whether public access to certain information is required to be delayed or limited, or accelerated.

Responses to specific Terms of Reference

Terms of Reference A

Report on the potential economic benefits and costs to Victoria of maximising access to and use of Government information for commercial and/or non-commercial purposes, including consideration of:

- i. public policy developments elsewhere in Australia and internationally; and*
- ii. the types of information that will provide the greatest potential benefit.*

And

¹ *VenturousAustralia – Building strength in innovation*, Review of the National Innovation System, 2008, p95.

Terms of Reference B

Consider whether the use of open source and open content licensing models, including Creative Commons, would enhance the discovery, access and use of Government Information.

Victoria will clearly wish to weigh the relative costs and benefits of restricting and increasing access to public information, including different ways of realising the commercial value of the information. However, CSIRO believes that Victoria could potentially derive significant economic and social benefits from maximising access to and use of government information in a number of areas. Some examples are discussed below.

1. Open access to geo-spatial data:

Currently Victorian geo-spatial data is provided to other government departments, but is not readily available to the research community. Detailed street and mapping data enables the development of new commercial products, such as GPS products, for the Victorian market. This has obvious economic benefits to the State through flow on production and consumption activity.

Open access of geo-spatial data to the research community could have social as well as economic benefits. For example state forestry data sets could be readily overlaid with CSIRO Forestry Division fire activity research to provide new insights into where and how bushfires are likely to spread within the State of Victoria and inter-State. Common access to the available data bushfires and bushfire conditions for both Victorian government and national researchers would have clear community benefits to the community and could contribute to preventing economic costs and loss of lives.

2. Open access to Victorian health data:

Free and open researcher access to de-identified Victorian public hospital data could provide a range of economic and social benefits to not Victoria, with flow on benefits for Australian and global public health. Increased access to volumes of data in relation to key diseases such as cancer, diabetes, hepatitis, Alzheimer's disease or HIV could expedite research into cures and allow for a broader range of simultaneous research to take place; specialists undertaking research would no longer need to be co-located with the hospitals, bringing a much broader range of possible analytical expertise to bear on the same data sets.

This is essentially the approach taken by the National Institute of Health (NIH) in the United States, which in 2003 mandated a "Data Sharing Policy" of ensuring open research access to (de-identified) public health data and records to enable more opportunistic and cross-disease analysis and cure identification. The policy objective was to enable all researchers, regardless of the size of their budgets, to have access to health research data acquired through publicly-funded activities (i.e. public health data). Both institutional and non-institutional researchers have responded positively to this development, which has opened up a shared information set and enable collaboration on a variety of levels to address national health research challenges.

Terms of Reference C

Report on the use of information and communication technology to support discovery, access and use of Government Information.

CSIRO is an active partner in the National Collaborative Research Infrastructure System (NCRIS). NCRIS 5.16 comprises the Platforms for Collaboration investment portion of

the entire NCRIS program, a \$75m investment in ICT infrastructure and services for research over three years. NCRIS investments will leave the research sector in Australia with a legacy in respect of:

- increased computational capacity through a national computation infrastructure (NCI);
- provision of collaborative toolsets and services through the Australian Research Collaborative Service (ARCS);
- access for researchers to a cross disciplinary research data repository service, enabling access to rich, managed sets of public research data through the Australian National Data Commons (ANDS);

The Australian Access Federation (AAF), an access and authentication service to enable collaboration via the sharing of data within and between higher education and research institutions both locally and internationally, as well as with other organisations.

The NCRIS program is working through the policy and practical issues associated with the provision of access to public sector information, including by establishing authentication services, and use and reuse policies. Victoria may wish to consider the possible benefits of possible benefits of engaging with this federal program, including

- Leveraging the \$75m commonwealth investment;
- Minimising the proliferation of different governance arrangements and data management regimes in Australia;
- Reducing inefficiencies in the current arrangements where publicly funded research organisation pay to access publicly funded data
- Increasing access to State data sets to support research directed to addressing national research challenges (egg the geo-spatial and water data sets mentioned above);

Terms of Reference D

Identify likely risks, impediments and restrictions to open content and open source licensing of Government information, including impacts on and implications for any existing cost recovery arrangements.

CSIRO observes that suitable governance processes will need to be in place concerning the publication of government information through open content and open source licensing models. In particular, prior to release of software or information through open content or open source licensing, it is recommended that a suitable clearance analysis be conducted to ensure that the government has the right to publish the information or to license the software or that publication will not attract inappropriate liability or loss of opportunity. Such an analysis would typically need to consider the following matters (to a greater or lesser degree depending on the context):

- Whether the code (in the case of software) or the information includes any confidential information of another party.
- In the case of software whether the code includes or embodies code prepared by another party (whether such third party code is obtained through open source arrangements or through another contractual arrangement with a party). The

terms under which such third party code was included need to be considered in assessing the government's right to release software that includes the third party code.

- Whether the information or software is subject to other contractual rights. For example where information such as research data is generated through a collaboration, the collaboration agreement may typically provide that certain processes are to be observed before the information may be published, for example to ensure that information is not published until a suitable decision making process has been completed concerning whether intellectual property protection should be sought, because premature publication will typically prejudice the capacity to obtain registered intellectual property protection for project outcomes including inventions.
- Whether the information is of a nature that could attract potential liability, for example through negligence, through constituting an infringement of third party intellectual property rights or non-compliance with trade practices legislation.

Putting in place suitable governance procedures to ensure that the above risks are properly assessed prior to release of information will require an investment in suitable personnel and infrastructure. It may be necessary to reflect the cost of such activities in making at least some categories of information publicly available.