



DRAFT

OSGeo AustNZ Logo?

**Open Source Geospatial Foundation
Australia - New Zealand Chapter**

**Submission
for
Parliament of Victoria's**

**Inquiry Into Improving Access To Victorian Public Sector Information
and Data**

Bruce Bannerman, <additional?>

22nd August 2008

©2008 Bruce Bannerman, other contributors?



This work is licensed under a Creative Commons Attribution 3.0 Unported License.¹

If you wish to reference this work, please attribute the authors and The Australia - New Zealand Local Chapter of the Open Source Geospatial Foundation.

Revision History

Revision	Date	Contributor	Comment
0.1	17Aug08	Bruce Bannerman	Initial Revision

¹ <http://creativecommons.org/licenses/by/3.0/>

Abstract

The Australia – New Zealand Chapter of the Open Source Geospatial Foundation (OSGeo-AustNZ) would like to congratulate the Parliament of Victoria for this exciting initiative.

We believe that one of the main impediments to the development of a vibrant spatial industry in Victoria is the difficulty experienced in obtaining free, quality Public Sector spatial Information.

This type of information is needed by the community at large for a wide number of reasons, including to help understand some of the Big Picture issues that are facing our society, for example Climate Change and Water Shortage.

This submission has been prepared by volunteer members of a new Open Source community.

OSGeo-AustNZ will be focussing on issues relating mainly to spatial information in its comments on PSI.

General comments

Answers to questions

Suggestions for community approach to maintenance of spatial data.

Submission may seem critical in places. It is hoped that this will be read as constructive criticism.

Key messages

Open access

Free data

Open Standards

Information Management

OS Maintenance

What is the Open Source Geospatial Foundation

The Open Source Geospatial Foundation, or OSGeo, is an international not-for-profit organisation whose mission is to support and promote the collaborative development of open geospatial technologies and data.

The foundation provides financial, organisational and legal support to the broader open source geospatial community. It also serves as an independent legal entity to which community members can contribute code, funding and other resources, secure in the knowledge that their contributions will be maintained for public benefit.

OSGeo also serves as an outreach and advocacy organisation for the open source geospatial community, and provides a common forum and shared infrastructure for improving cross-project collaboration.

OSGeo has attracted a large on-line international community who work collaboratively to further the cause of open source and open data as it relates to spatial information.

More information on OSGeo can be found at:

<http://www.osgeo.org/content/foundation/about.html>

What is the Open Source Geospatial Foundation – Australia New Zealand Local Chapter

The Australia – New Zealand local chapter of the Open Source Geospatial Foundation (OSGeo-AustNZ) aims to act as OSGeo's local representative within the Australian and New Zealand region, furthering the aims of the Foundation.

OSGeo-AustNZ is in the process of incorporating as a Not For Profit organisation in Australia.

In its short time in existence, OSGeo-AustNZ has attracted a significant community, including a number of key people in the regional spatial industry.

The community is proving to be an effective way to break down the traditional barriers that exist between different levels of government, research organisations and private enterprise. It is also encouraging to see an increasing degree of cooperation and cross fertilisation between Australian and New Zealand organisations.

OSGeo-AustNZ is contributing to this inquiry in line with one of our organisation's key goals:

- To promote freely available geodata - free software is useless without data.

More information on OSGeo-AustNZ can be found at:

<http://wiki.osgeo.org/wiki/Aust-NZ>

The main contributors to this document on behalf of OSGeo-AustNZ have been:

- Bruce Bannerman

-

A summary of contributor's backgrounds may be found at Annex A.

General Comments on the Enquiry

OSGeo-AustNZ congratulates the Parliament of Victoria for taking the courageous step calling this enquiry.

We believe that this is a timely enquiry and call for radical and sweeping changes to open up access to PSI.

What are some of the current problems being faced trying to access public sector spatial data in Victoria?

There are a range of issues that affect access to public sector spatial data in Victoria. These include:

Intellectual Property Issues

In addition to the Intellectual Property issues that have been covered in the Inquiry Discussion paper, there is also the issue of what IP applies to spatial data that is purchased on behalf of the government. Some examples are:

- Aerial photography;
- Satellite imagery; and
- formatted vector data.

There have been a number of problems experienced over the years that could be quite difficult to resolve, and these problems should be resolved because a lot of the data in question may have relevance in the future to help resolve Climate Change issues.

Some examples of problems are:

- Aerial Photography was tendered for and purchased by a government department under an agreement that ceded IP to the vendor, as the cost was 'cheaper'. Unfortunately, no other government organisation was, or is, able to use this valuable dataset due to the restrictive IP. The vendor is happy to sell the government additional copies of this data though.
- Similarly, some of the high resolution satellite vendors have very restrictive arrangements where they retain the IP in the data and constrain what can be done by purchasers.
- There is also a trend among some commercial vendors to offer to sell attractive datasets to government. However, if the license agreement is looked at there are restrictive usage arrangements in place such as 'the data may not be used in an externally facing web site'. This renders the data unusable in most instances.
- In a recent High Court of Australia judgement against NSW, the judgement found to the effect that the NSW Government was not entitled to copy survey plans and provide them to the public without providing remuneration to the copyright holders. This has broader implications for the provision of PSI to the public.
- There is an increasing trend towards 'seamless' spatial datasets that cover the geographic

extents of the State of Victoria. These datasets may comprise spatial data that has originated from a variety of sources with differing IP arrangements in place. The people who construct and use these datasets may not be aware of IP related issues, or of what part of each dataset came from which source data.

Cost Issues

There are anecdotal reports that it can be quite costly to access spatial data in Victoria.

Internally to government, there have been reports of departments who have been asked for a substantial annual fee per user in order to access spatial data. This is for an access fee to spatial data that the government has already purchased or captured internally.

External to government there are also reports of substantial fees being asked to access spatial data. (we need examples)

In part this issue is driven by government budgetary issues. A substantial portion of Spatial Information Infrastructure's (SII) budget for example comes from the sale of 'licenses' to spatial information. SII is a Division within the Department of Sustainability and Environment.

In part, this issue is also driven by the need to try and secure some (inadequate) funding to contribute towards spatial data maintenance. Maintenance will be covered separately, below.

Information Management Issues

Victoria is facing an increasing number of challenging issues that demand ready access to quality spatial information. Several of these issues are Big Picture issues that are going to require innovative and possibly even radical approaches to try and resolve. Some examples are issues such as Climate Change and Water Shortage.

Government does not have a monopoly on expertise relevant to these areas. However, Government is certainly a key partner in trying to help find solutions. As a good partner, one of the items that Government can offer the community to help resolve these problems is free and unrestricted access to the Government's spatial data holdings.

To do this, Government needs to have good Information Management processes in place to make it easy to search for, find and deliver data that is applicable for a particular purpose. Outlined below are a number of perceived problem areas that may make it difficult for the Government to achieve its aim of Improving Access to Victorian Public Sector Information and Data.

Data Custodians

Spatial Information Infrastructure (SII) have prepared some excellent strategic guidelines for spatial data custodians. The aim is to get accountability for spatial data sets to defined people.

There are a number of conscientious spatial professionals in a number of departments who are attempting to implement these guidelines, however, they do not appear to be adequately supported by their senior management and as such, the guidelines do not appear to be widely adopted.

If the public sector is to retain Custodianship of defined spatial data sets, Custodianship must be allocated at a senior management level and monitored via the performance management process. The Senior Managers should then be help accountable for the availability and quality of that data.

It is understood that an attempt has been made in this regard, however judging by the funds that are available for spatial data maintenance and other Spatial Data Infrastructure issues, this does

not appear to have been successful.

Spatial Metadata

Spatial Metadata is structured information that describes a given spatial dataset, including properties such as: search key words, data quality, intended use, access constraints, contact people, physical location of the dataset and the spatial area that it represents.

Spatial Metadata is intended for use within a Spatial Metadata Catalogue to help other spatial professionals find spatial datasets that could potentially help them with a particular piece of analysis or other task that they need to undertake.

Ideally, a professional could save a considerable amount of work and money by browsing a catalogue looking for existing data that may be suitable for their purpose.

Unfortunately, metadata and metadata catalogues are not given appropriate priority by managers, data custodians and spatial data maintainers. Some examples of problems are:

- Metadata may be entered once, but it is rarely maintained.
- Metadata may be entered for key datasets, but there is a lot of other spatial data created as part of a project, or maintained in traditional business systems that do not necessarily have metadata records.
- Australia is in the process of migrating from an old metadata standard to an international standard. Unfortunately, the Victorian corporate project to implement a system and catalogue to support this standard has not been adequately resourced and is consequently taking a considerable amount of time to implement.

If the Government does not improve its focus and resourcing on issues relating to Metadata, it will be very difficult for the public to be able to find public sector spatial information.

On a positive note, SII have several people who are well respected within the Australian Spatial Community and have contributed significantly to the implementation of the new metadata standard.

Project Focus and Cost of Disk Space

While Victoria is developing a number of reasonable contiguous statewide spatial datasets e.g. the VicMap series of products and the Seamless Geology dataset, there is a class of spatial data sets that are often overlooked.

Often spatial professionals will be asked to undertake some analysis or research work within the context of a project within their business unit. Often the output of this work will be a written report with a set of printed maps and a presentation.

What happens to the valuable spatial data that was created during this analysis or project work? Some of this data may well be of relevance to a large number of people.

Often this data is archived at the end of a project or possibly even forgotten about as people move on to other projects. Metadata may or may not be created to describe the data created.

There are no clear processes in place to help a project team to decide whether the data that they have created is of corporate interest or not.

Some preliminary analysis work that was undertaken during the Gold Undercover Project for the Department of Primary Industries highlights a couple of areas of concern and may well be indicative of what is happening in other departments as well. It was found that Spatial data tended

to be archived to individual users desktop PCs, Portable Disk Drives, or to tape without adequate backup.

There were a number of reasons for this, including:

- The department did not have a corporate spatial data maintenance environment. It is understood that this is the normal situation in most departments.
- The cost of enterprise storage solutions available is very expensive within government. The rate quoted internally to host 1TB of data per year is nearly 20 times more than that quoted by a well known commercial service. Therefore project personnel opted to archive their project data on Portable Disk Drives or on tape.
- The department has approximately 60 TB of spatial data. Given the costing issue, of this, over 55 TB is stored off site on tape. Given that tape has a limited life span, this data is at serious risk.

These issues will also need to be addressed to improve access to Victorian PSI.

As it currently stands, significant holdings of data are not available for use, search, or delivery.

Data Format Issues

Proprietary solutions and proprietary data formats

Within the Victorian Government, there is an unhealthy focus on the use of proprietary solutions and proprietary data formats for managing and delivering spatial PSI.

This has led to a range of issues that affect the Government's ability to improve access to PSI. Some examples are:

- Focussing on proprietary vendor solutions has led to vendor 'lock-in' in some situations. One example of the problems that this can cause government is:
 - Ten to fifteen years ago, one particular spatial vendor had a range of products that were well regarded with the state government. Individual branches and departments customised some of these products considerably, integrating them with their business systems to deliver advanced functionality to their staff and the public. An example is the Minerals and Petroleum Division's Eureka! application.
 - This vendor introduced a new product line, but did not provide an upgrade path for the the superseded products. This has left the Government with no upgrade path for such applications and the prospect of finding sufficient funds from budget bids to redevelop the same functionality, using newer technology.
 - Consequently as of 2008, many of these business customisations have still not been replaced. The applications that they were based on, are well past their end of life status with no support from vendors. Due to the perceived archaic nature of the technology, it is also very difficult to find suitably trained personnel to keep them maintained.
- Providing data in specific proprietary formats places constraints, undue expense and additional work on people and organisations that wish to consume the data. Some examples are:
 - To fully utilise a spatial dataset as it was intended, requires the use of the data in it's

native data format (excluding open standards, see below). If the data is not used in this format and translated into another format, then there is often contextual information that is lost during the translation (e.g. topology, relationship rules etc).

- Post processing is often required by people translating proprietary spatial data to get it into a format that is suitable for their use. Each time that the source data is updated by the Government, which in some cases occurs on a regular basis, the people importing the data must go through this same process again.
- Consumers of Government spatial data may be forced into buying very expensive software in order to utilise Government spatial PSI so that they can minimise the amount of rework that they have to do in converting proprietary data formats.
- The use of proprietary spatial data formats for the **Archival** of spatial PSI is also a high risk strategy. Given the Government's experiences with vendor lock-in and lack of upgrade options, it is highly likely that in ten years time that spatial PSI that has already been archived in proprietary formats will be unreadable. See also comments on Open Data formats below.

Spatial Imagery data formats

This issue is also related to the cost of disk space issue as discussed above.

Victoria has a very rich holding of spatial imagery that spans a number of decades. This imagery will prove to be an invaluable data source to help address current Big Picture issues such as Climate Change and Water Shortage.

Traditionally, spatial imagery has been treated as a 'pretty picture' and utilised effectively in mediums such as Google Earth.

What is often missed is the value that this data offers as an input into digital image analytical processes. Such techniques can be an invaluable input into spatial analysis projects.

In an example relating to Water Shortage, a multi-temporal analysis could potentially be conducted over a particular area to determine the difference in surface water between e.g. 1975 and 2005 (assuming that source data is available).

The long term ability for Government, academia, industry and the public to undertake such analysis in the long term is at risk.

There is currently a view within a number of organisations in Government that the majority of users of spatial imagery only want to view imagery as a pretty picture. Therefore, current plans are that this imagery will be compressed into a particular format and source imagery backed up to tape.

The problem with this approach is that:

- People who want to undertake analysis, will wish to utilise the source data, not the compressed data. This will not be readily available. Once the data is on tape, there is a very real risk that it may not be able to be recovered at all due to corrupted media.
- Given the invaluable nature of this data for 'Big Picture' analysis, this is a high risk strategy.
- There is a very real doubt that the radiometric integrity of the imagery compressed in the format proposed for storage could be restored in the future in the event that the original data could not be recovered from tape.

Open Standards – Open Geospatial Consortium and ISO 19100 series

There is a rapidly increasing demand in the world for the use of Open Standards for both the design of IT solutions and for the provision of data.

Using the vendor lock-in example above, if these systems were to be designed today, it would make sense to design them around the use of open standards to insulate the systems from any changes to vendor technology. If designed appropriately, the use of open standards should allow one vendor's product to be replaced by another, avoiding the type of problem that is currently experienced necessitating the same functionality to be redeveloped and essentially re-paid for.

Similarly, the use of open standards for the provision of data is having a revolutionary affect on the provision and consumption of spatial data. As long as producer and consumer both have products that understand how to utilise a particular set of standards they can utilise data with no translation and consequent rework and loss contextual information. See also comments on GeoSciML below.

The two most relevant sets of open standards relating to spatial data are:

- Open Geospatial Consortium (OGC):
 - OGC Standards have been in ongoing development and refinement for over sixteen years. They are built on a solid theoretical base.
 - They are rapidly gaining market acceptance, and seen as mechanism that organisations can use to seize back strategic control of their IT investments.
 - Open Source spatial applications, including OSGeo applications, have very strong support for OGC standards. Most of the OGC standards specify an Open Source product as the reference implementation for that standard. These products are available free of charge to end users.
 - One of the OGC standards, GML has strong potential as an archival medium for vector spatial data.
 - It is also fair to say that OGC standards are evolving as the spatial industry's understanding of its requirements for open standards are also evolving. Therefore, it would be prudent for the Victorian Government to become actively involved in the development and implementation of these standards. Australia has already contributed a significant amount to the development of these standards. For example, Dr Simon Cox from CSIRO was awarded OGC's prestigious Gardel's Award for his outstanding contributions².
- ISO 19100 series of standards
 - The International Standards Organisation has a set of standards that relate to spatial information.
 - These standards are directly relevant to Australia.
 - ISO and OGC work closely together on the development of spatial open standards.

Open Standards – GeoSciML a Victorian Government example

Over the last five years, a group of Victorian Government employees from the Minerals and Petroleum Division have been helping to pioneer the development of spatial open standards pertinent to geoscience data.

² <http://www.opengeospatial.org/pressroom/pressreleases/577>

They have been working with a group of like minded professionals from a number of key geoscience organisations around the world to develop ways to exchange and utilise geoscience information via the use of spatial open standards (OGC standards). This team has pioneered the development of GeoSciML³.

The world-wide GeoSciML team have had a significant achievement. They have been able to get professionals from a number of countries, speaking a range of languages to agree to a common way of describing and exchanging data representing one of the most abstract types of spatial information available.

Using this open standards format, it is now possible to take geoscience data served by disparate organisations from around the world and combine this data into a single homogeneous dataset that can be analysed as a single entity.

This team is showing what could be possible to help society develop profiles of a range of data types to help specialists understand factors relating to Big Picture issues e.g. Climate Change and Water Shortage.

Business Context Issues

Traditionally, spatial data has been seen as an input whose sole purpose has been to produce a map. Much of the spatial data delivery processes used within the main delivery mechanisms, the Corporate Spatial Data Library (CSDL) and Spatial Data Mart (SDM) tend to reflect this concept.

Increasingly spatial information is seen as a critical decision support tool. Often analysis is conducted that requires a set of spatial data and a number of aspatial data sets from business systems.

There may be a direct relationship between the spatial and aspatial data in the context of a primary / foreign key relationship. However, equally as relevant there may only be a spatial relationship between the two, e.g. all properties that are within 100 meters of bus route 790 to Frankston.

Both the CSDL and SDM do not cater for live links to aspatial data in business systems. They rely on a spatial data maintenance person manually creating a spatial dataset with a subset of the appropriate attributes (if possible) and copying this into the CSDL for consumption by the SDM.

This is not suitable for:

- Time sensitive data such as mining tenements.
- Dynamic exploration of data relationships.
- Complex data such as GeoSciML.

Spatial Data Maintenance Issues

Spatial data requires ongoing maintenance, if it is to be utilised effectively as an aid to decision making.

There is an argument to the effect that:

- The US provides its spatial data for free, yet considerable portions of its data are not well maintained.
- In the UK, the Ordnance Survey charges for access to its data and it has a well maintained

³ <http://www.geosciml.org/>

library of spatial data.

- Therefore, if you charge you your data you can have well maintained data.

This argument misses a few key points:

- It does not take into account the geographic size differences between the USA and the UK, assuming that it costs the same to maintain data for both countries.
- Victoria has approximately the same geographic size as the UK and similar policies in charging for access to spatial data as the UK. Yet as highlighted by the Nationals Member for Benalla⁴ Mr Bill Sykes, there are maintenance problems with Victorian spatial PSI.

It can easily be argued that the types of problems identified by Mr Sykes are symptomatic of decades of neglect in spatial data maintenance.

Much of the VicMap data is still believed to be based on 1970s data.

Blame for this state of affairs cannot be placed on SII or its predecessors. They have been doing an admirable job in an environment of successive budget cuts.

Blame for the current state of affairs must lay with successive Parliaments for not ensuring that sufficient resources are provided to ensure that such a critical resource is adequately maintained.

OSGeo-AustNZ has a proposal for spatial data maintenance at the end of this proposal that may result in win-win situation for Government and industry.

Spatial Data Liability Issues

There is a great deal of uncertainty over liability relating to the use or misuse of spatial information.

More required here.

4 <http://www.abc.net.au/news/stories/2007/12/06/2111673.htm>

Comments on Chapter 1

Question 1: What are the advantages and disadvantages of government adopting 'push' and 'pull' models toward the publication of public sector information (PSI), respectively?

Comments on Chapter 2

Question 2: How can improved access to and re-use of PSI drive economic growth, employment opportunities and new commercial ventures?

Free and unrestricted access to spatial PSI can have a significant impact on economic growth, employment, opportunities and new commercial ventures. Some examples are:

- Quality spatial information is a pre-requisite planning aid for many engineering and infrastructure related projects.
- This information can have a significant impact on the quality of early analysis and on investigation of alternate what-if scenarios.
- By providing this information for free and with unrestricted access, the Government can reduce preliminary planning costs considerably. This in turn could lead to more money being available for other engineering work and potentially more projects.
- This in turn would lead to more activity and employment within dependent businesses and communities. There would also be money returned to Government by way of taxes, royalties etc.

There is also another challenge and potential benefit that could arise from this approach. Engineering and Infrastructure related organisations spend considerable sums of money on more detailed spatial information at later stages in project life cycles. The challenge will be to foster a community attitude where it is normal to exchange spatial data for free and with no constraints.

Question 3: What can the Victorian Government do to improve access to PSI in a manner that creates new opportunities for information and knowledge flow, and thereby encourage further innovation?

The Victorian Government can provide PSI for free, without restrictions⁵ via open standards. In the context of spatial PSI, it is assumed that this information would be provided via OGC Web Services.

Having such data available via OGC Services would encourage innovation via the use of Web 2.0 techniques such as Mashups.

Question 4: If the Victorian public sector is to provide increased access to information, what kind of information would provide the greatest opportunities to improve or develop ?

All spatial information should be provided for free, without restrictions via open standards.

This includes, vector, grid, related aspatial and spatial imagery.

The term 'spatial imagery' is used in its broadest context throughout this document and includes all kinds of remotely sensed data including for example: aerial photography, satellite imagery, radar, lidar, gravity, aeromagnetism, multi and hyper spectral, sonar etc.

It is important to understand that historic imagery is just as important as current data.

⁵ OSGeo-AustNZ acknowledges that there would need to be some constraints on data access, e.g. to comply with Privacy Legislation.

Question 5: How can social engagement, in particular through the development of spontaneous social networks, be enhanced through the provision of enhanced access to PSI?

The provision of PSI for free, without restrictions via open standards can help to drive the spontaneous development of social networks.

This can happen in a manner similar to the way that OSGeo-AustNZ itself has evolved, for example:

- After cursory exploration of Victorian Government PSI, several people see the potential for a much deeper analysis. They may be looking at a particular aspect of one of the Big Picture issues.
- They establish a mailing list and announce their intent to mailing lists where they know they will find people with similar interests.
- In time, other people will join their list and they will begin to develop a 'community'.
- Eventually, they will attract sufficient people to be self sustaining. Like many such communities, there are no borders, so they may well attract the interests of world leading experts from Australia and overseas.
- One of the attractions binding this community together would be the ready access to quality Victorian Government PSI that is available for free, without restrictions via open standards.
- The community would use this data to explore the issue in question.

Question 6: In what circumstances can open access to PSI empower individual citizens and communities to participate in social and political activities?

In a manner similar to that described in the answer to question 5.

Comments on Chapter 3

Question 7:

.

Annex A - Summary of Contributors

Bruce Bannerman

Bruce is a Spatial Information specialist with approximately thirty years experience in the industry. He has worked around Australia in Local, State and Federal Government. He also has substantial experience in private enterprise, large and small, including a period of five years as an entrepreneur, operating a small spatial consultancy.

Bruce is an advocate for Open Source software and free and open access to spatial data.

He has experiences that are relevant to this enquiry, including:

- For the past two years, working as an IT Solutions Architect within the Department of Primary Industries looking at issues relating to opening up Departmental data for Search and access by the external public.
- As a Director of OSGeo-AustNZ.
- As a Director of a small spatial entrepreneurial consultancy, GeolInnovations Pty Ltd (currently not trading).
- As a member of the Victorian Spatial Council's Coordinated Image Acquisition Work group.
- Personal interests over the last two decades in spatial data management and access issues.