



*SUBMISSION TO THE*  
*INQUIRY INTO IMPROVING ACCESS TO VICTORIAN GOVERNMENT PUBLIC SECTOR*  
*INFORMATION AND DATA*  
*INTELLECTUAL PROPERTY RESEARCH INSTITUTE OF AUSTRALIA*

**Our organisation**

The Intellectual Property Research Institute of Australia is a multi-disciplinary research body of the University of Melbourne. It includes 10 research fellows from economics, law and management and over a dozen research associates. Its mission is to conduct and disseminate research into the economics, management and law of intellectual property (IP).

IPRIA's research focuses on ways to improve the protection, management and exploitation of intellectual property by business, research institutions and other users of the IP system, and on supporting high quality policy development by government in areas relating to intellectual property. It seeks to use the outcomes of its research to create and contribute to healthy public debate on key issues relating to intellectual property.

This Submission was prepared by Associate Professor Elizabeth Webster and Dr Chris Dent.

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

The provision of information is a classic example of a market failure. The optimal provision of information and data requires that information and data be priced at its marginal cost. In many cases, this implies a zero price because the incremental cost of an extra user is negligible. However, no for-profit company will begin to produce (collate, clean, document and distribute) information and data at a zero price. Accordingly, some type of government intervention is required to furnish the incentive to collect



information and data that is used by others.

One type of government intervention is to create IP rights. Conferring IP rights upon the owner of the information and data may solve this incentive to produce problem by enabling the owner to exclude people from consuming the information and data. Exclusion empowers the owner to charge a non-zero price. If this price is high enough to cover the (substantial) average fixed or one-off costs of production, then there is a clear incentive to collect the data.

However, while a price ‘solves’ the incentive problem, it creates a deadweight problem for classes of goods (and services) called non-rival goods. Putting a price on a product is only justified if by denying some consumers from using that product (because they will not pay the price) we allow the product to be consumed by others who value it more or we allow the resources to be re-directed to more valuable products. A classic example where a price has a positive role in determining efficient resource allocation is consumer goods. If the price reflects the marginal cost of producing the product, then a consumer will only buy it if they value it at least as much as the resources that are diverted away from other uses. However, information and data are special products and do not conform to this classic case. Information and data are non-rival goods which means that one person’s use of a piece of information does not affect another’s ability to use that information. All potential users who value the information and data by less than going price will not use the information and data even though its use will create social benefits with no attendant social costs. The loss of these net benefits is called a deadweight loss.

If we want to provide the optimal level of information and data, the fixed costs of its collection, cleaning and documentation should be paid through the public purse, but the data should be ‘sold’ at a price equal to its marginal cost. The latter includes the additional costs of bespoke requests.

Note that a ‘price’ does not need to be a monetary value. It can be the time spent by the customer working their way through complicated and hostile agencies, or having to comply with complex legal contracts.





The more simple and straightforward the means of obtaining the information and data, the lower the 'price'.

Net benefits are reflected in a more efficient production chain and higher productivity.

**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: a) investment and business opportunities? b) social, medical and scientific research? c) community and civic engagement**

Information and data add to social welfare by:

- Enabling businesses to make more accurate decisions about what to produce and where to produce it. This will reduce costs of production and enhance competition.
- Enabling businesses to make better market forecasts and thus reduce the waste associated with errors in production and marketing,
- Enabling community groups to more accurately monitor the welfare of their interest group
- Providing researchers with more accurate information on which to base their evidence-based policy advice.

We do not know of any empirical studies which have tried to assess the social welfare losses when public sector information and data is 'sold' at prices that deny access (a price may be institutional resistance rather than a monetary value). However, it is our experience that lack of access to information and data from government and public sector archives is a frequent block to research. This frequency depends on the attitude of the actual agency. Some government and public sector bodies recognise that they hold a unique and valuable resource and make internal resources available to prepare the information and data so it can be used by outside parties. Others actively resist requests for data. Many however, merely allocate no staff time to the task (and do not consider a cost-recovery policy) with the result that only the very persistent or well-connected outsider can get access to the data.





It is our experience that it is the culture of the agency, and therefore the decision by the CEO to allocate staff time the task, which determines their willingness to provide information and data to outsiders rather than the nature of the information.

The potential users of information and data are wide and varied and the uses are likely to be substantial. Examples include:

- Businesses who want information on likely market demand;
- Businesses who want locational information on competitors
- Community groups who want objective data on the magnitude of the problems they face
- Researchers who want to assess the size and scope of social and economic issues
- Researchers who want to objectively assess various policy options.

**Question 7: What institutions and agencies should be considered part of the public sector for the purposes of this Inquiry? What advantages will be obtained by encompassing some or all of the following agencies and institutions under this definition: a) executive government: principally government departments, but also incorporating statutory authorities? b) the legislature: including parliament? c) the judiciary? d) local councils? e) other public institutions, such as universities, TAFEs, public hospitals, etc?**

There is no reason why the meaning of public sector should exclude local councils, hospitals, universities, colleges, semi-government agencies and statutory authorities. Some bodies, such as hospitals, clearly hold more sensitive data collections than others and this should be respected.

However, special provision should be made for commercial units within these bodies. Commercial units may already be collecting and selling data (at full cost recovery). Forcing these units to sell at marginal cost will force their closure. Examples include economic forecasting which is performed in academic research centres, drug trial data which is performed on behalf of pharmaceutical companies. Consideration should be made to including private sector agencies which have the monopoly rights over the collection



of lists such as telephone lists.

The Committee's attention may be drawn to the Victorian Law Reform Commission's Report, *Criminal Liability for Workplace Death and Serious Injury in the Public Sector, 2002*. In that Report, the Commission considered the application of the then proposed *Crimes (Workplace Death and Serious Injuries) Bill 2001* to the public sector. Many of the issues that arise with respect to identifying the Crown and public sector for the purposes of the Bill arise with respect to the legal identification of the Crown and the public sector for the purposes of Crown copyright and the application of processes for the publication of public sector information.

**Question 8: What kinds of documents, data and/or other materials should be considered for public access? What criteria should be applied when judging whether specific documents, data and materials should be made available to the public?**

Release of information and data should however be limited by the needs of:

- Commercial confidentiality. Information and data should not be released in a way that transgresses commercial confidence.
- Personal privacy. Information and data should not identify individuals.
- The costs of cleaning and documenting the information and data. The theory described above assumes, for pedagogical reasons, that marginal costs are (near) zero. In practice this is not the case and there can be non-trivial costs of cleaning, documenting and distributing the data.
- The commercial basis (as a for-profit activity) of the collection. If the data was collected as a for-profit activity, then the price charged to consumers needs to reflect the full costs of collection and preparation. If however, the collection was funded through taxes (and is therefore a not-for-profit activity) then the price should reflect only the costs of making the data usable (i.e. cleaning and preparation). Some public sector agencies – such as government business enterprises, universities and local government authorities – have commercial units located within them.

Much commercial information and data are not confidential but are available through on-line sources (such as company reports, ASIC documents, law court documents). However, these data are not available in a



conventional database format and are thereby not readily accessible for businesses, community groups and researchers.

Aggregation is the main way commercial data can be confidentialised. However, aggregated data is usually only useful for descriptive purposes and is often not useful for research or analytic purposes. Researchers generally require access to micro (unit record) data to conduct rigorous research.

Access to unit record data for researchers can be made without breaching confidentiality by:

- Removing individual identifiers (of companies) from the data (such as specific location, company name, date of birth ), and/or
- Allowing access by vetted individuals at the data collection site, and/or
- Requiring researchers to sign confidentiality clauses.

The level of limitation would depend on the sensitivity of the dataset and the range of people who would want to apply for data access.

Overseas national statistics agencies, such as Norway, currently operate best practice in the provision of commercially sensitive company data to researchers in unit record format.

In no circumstances should there be a blanket obligation to release the original data as this would contravene conditions under which the data was collected. Releasing the original data would be unethical and would ultimately cause many individuals and companies who supply the data to the public sector agency to refuse to supply the data in the future.

**Question 11: What criteria should government apply when determining whether to provide access to PSI? Under what circumstances would the following pricing options be appropriate: a) no cost? b) marginal cost or cash recovery? c) commercial profit and return?**

As stated above, if we want to provide the optimal level of information and data (which minimises



deadweight losses), the fixed costs of its collection, cleaning and documentation should be paid through the public purse, but it should be ‘sold’ at a price equal to its marginal cost. How large this ‘margin’ is depends on whether we define the marginal consumer to be all outside users of the information and data, or, one extra outside person. If it’s the former, then the marginal costs probably include much of the cleaning and documentation costs. If it’s the latter then it is close to zero.

It is reasonable that public sector agencies can expect some level of cost recovery from users for the extra costs of cleaning and documenting the data – that is the cleaning and documentation that is over and above what they do for internal use. Only agencies which collect the information and data through their commercial arms should be allowed to recover part of the collation costs through its pricing strategy.

Universities collect primary data for research purposes. It is common for these to be only partly funded through Australian Research Council grants (grants are rarely made at 100%). Accordingly, it is not possible to know what portion of the costs of collation has been funded by the public purse and what has been funded through the researchers free time. Some provision in the rules should be made to ensure that the person who collected the data using non-government funds is protected from being forced to release the data to whom ever applies. If some level of protection does not exist, there is likely to be a negative impact of data collection for research purposes. This includes data across all disciplines.

### **Question 12: What other open content licensing models may be of interest to the Committee?**

IPRIA, in its previous research,<sup>1</sup> has looked at only one model for sharing copyrighted material that (a) is not included in the Committee’s Discussion Paper; and (b) is not reliant on the Creative Commons (CC) suite of licences.<sup>1</sup> That model governs the distribution of educational content by the Learning Federation – an initiative of Australian state, territory and federal governments and the government of New Zealand.<sup>1</sup>



The licences used by the Learning Federation to distribute the material are categorised in terms of the nature of the organisation that is seeking to use the material. Organisations which licence material include schools (both government and private); universities and other bodies that train teachers; cultural institutions; and ‘vendors’ who seek to use the material for ‘testing and demonstration purposes’.<sup>1</sup> One notable limitation of the “vendor” licences under this scheme is that they do not include rights relating to the communication of the material or to the creation of derivative works. There are jurisdictional limits to the licences granted by the Learning Federation.

**Question 13: Is the absence of conditions regarding geographical restrictions or no endorsement in Creative Commons likely to be an issue for PSI?**

The lack of capacity to modify, on a case by case basis, the terms of CC licences may be of concern. It has been suggested, for example, that a significant reason for including a geographical restriction is that it may be seen to facilitate enforcement of the licence. A further reason is that included in the discussion paper – if Victorian, or Australian, public funds contributed to the production of the information, why should those outside the jurisdiction benefit? A response to this argument is that an open access policy here could encourage an open access policy in other jurisdictions; thereby making available to the Australian public a much wider range of materials than if all countries adopted a restrictive approach.

The question of any perceived need for a “no endorsement” term could turn on the nature of the information subject to the licence. Public sector organisations tend to produce a wide range of materials. It may well be problematic for the government in circumstances where the re-use of some of these materials by a licensee implies an endorsement of the goods, or activities, of the licensee. The reproduction of government nutrition guidelines in conjunction with an advertisement for a health food could, for example, encourage the public to infer that the government endorsed the food product. Issues relating to endorsement, however, may be less of a concern, for example, with respect to the reproduction of legislation.





**Question 14: What are the merits of the Victorian Government developing its own whole-of-government licensing framework as an alternative to adopting the Creative Commons licensing system?**

IPRIA's research into open content licences concluded that it would not be feasible to adopt the CC licensing framework as the sole framework for governing the distribution of PSI. This is because the simplicity of the CC framework does not adequately match the needs of the release of PSI. In other words, the most significant advantage to developing a licensing framework separate to the CC licences is that it can be tailored to the needs of the government. Further, in some cases the documents to be released may include copyrighted material from a third party – that third party may not agree to licence that material under any of the available CC licences.

One particular sets of circumstances emphasises the need for licence terms that do are not part of the CC suite of licences. There is no reason why the products created from the use of government data should be forced to the made publicly available. Many of the users of public sector information and data are tax-paying for-profit companies and it is in the national interest for them to use data to improve the efficiency of their production system (planning, marketing, forecasting etc). Demanding that they make their end-use products openly available to all others will effectively preclude their using the original information and data for commercial purposes.

A similar argument can be mounted for governments who want to commission private research to enable them to improve their policies. If the commissioned research has to be made public, then the research will in many cases not be commissioned. Forcing too much public disclosure merely drives people underground.

It is in the national interest therefore that the users of public sector information and data – whether they are businesses or government departments – have the option over whether or not they disclose the outcomes of their use of this data. We do not support the use of open content licenses that do not provide an option of





non-disclosure in appropriate circumstances.

A single whole-of-government framework would also enable a state government “brand” of licences – a brand that could include standard licences and open content licences. An assessment of the importance of such a brand is outside the scope of this submission, however, there may be efficiencies to a simple process for negotiating all licences for PSI. This process could be conducted by a centralised body – this could also lead to government-wide consistency with respect to copyright management.

There are, nonetheless, costs associated with a whole-of-government framework. There would be costs associated with establishing a set of tailored licences and there would also be costs associated with maintaining the framework. These costs are likely to be in excess of the costs associated with adopting CC licences across the public sector. It may be possible for the Victorian government to work with the federal and other state governments to develop an Australia-wide public sector open content licence – this could lead to reduced costs to the Victorian government in terms of the establishment of a tailored licensing framework.

**Question 15: Is it appropriate for the Victorian Government’s licensing framework to comprise both the Creative Commons licences and other more tailored licences?**

There are advantages to the simplicity of adopting the CC suite of licences, however, given the range of material that is potentially available from the public sector, there could also be advantages to having more tailored licences for some of the material. The simplicity of the CC licences, in part, stems from the four variables that form the basis of the licences – the nature of the use; the potential requirement for attribution; the capacity for derivative works to be produced and the terms that attach to the distribution of derivative works. CC licences, therefore, may be suitable for some PSI, such as guidelines that are aimed at wide distribution. Our research does not reveal any evidence that CC is not appropriate for PSI, however, there is also little evidence that the use of CC licences is effective in the public sector – though this is due



to the relatively short time frame that CC licences have been used by the public sector. For material that is not suitable for release under a CC licence, then, more tailored open content licences may be appropriate. If an agency had PSI that it could release but the information is not suited to a CC licence, then if there is no tailored licence available, the agency will not, and possibly should not, release the information. The withholding of the information may not be in the public interest.

**Question 17: What are the range of licence conditions that the Victorian Government is likely to require when issuing open content licences?**

As detailed in the above responses, the conditions that attach to CC licences may be suitable for some Victorian PSI. Additional requirements that could be included in tailored licences are:

- Restrictions with respect to the purposes the material could be put – such as for educational purposes only (such as exist in BC Commons licences) or there could be ban on derogatory use (such as in the BBC Creative Archive licences);
- Geographical restrictions;
- No endorsement terms; and
- Clauses that require accuracy in any reproduction of the licensed material.

Each of these would provide greater confidence that information produced through the expenditure of public funds would not be misused.

A further licence condition could be a capacity for the government to change licensing terms (with notice to the licensee) at a later date. Such a condition would facilitate greater control over publicly funded information.

Thursday, August 28, 2008

