
Access to Online Legislation and e-Government

Michael Rubacki

**Presentation at Southern Currents: Joint Conference ALLG &
NZLLA, Melbourne, September 2006**

27/9

Table of Contents

1	Overview	2
1.1	Overview	2
2	Brief history of online access	2
2.1	Early history from 1990	2
2.2	Mid history	3
2.3	Recent history	3
3	Ownership and policy underpinnings	4
4	Ownership & control examples	5
4.1	S. E. Asian illustration	5
4.2	Local illustrations	5
5	Technology	6
5.1	Word processing to electronic publishing & XML	6
5.2	Data capture and markup	6
5.3	Metadata	7
5.4	XML systems: requirements and overheads	7
5.5	Scanning: accuracy and quality control, NSW experience	8
5.6	Scanning older material	8
5.7	Future capture	9
6	Paper products	10
6.1	Static, unchanging range	10
6.2	Limitations, numbers and costs	10
6.3	Future?	10
7	Online products	10
7.1	Infancy	10
7.2	New products	11
7.3	Integration with legislative process	11
7.4	Development path	11
8	Collaboration	12
8.1	Record so far	12
8.2	Law librarians	12

1 Overview

1.1 Overview

Drafting offices and Law Librarians should intersect more often – Tasmania's Dorothy Shea addressed the annual PCO ITF in 2005 and I was very pleased to be invited here today. Let's hope that this becomes a more frequent arrangement.

Viewpoint. This presentation is based on events of the last 15 years but is very much a personal viewpoint. It does not necessarily coincide with the policy views of the government in NSW or the other jurisdictions mentioned.

My perspective is that of someone who believes that legislation should be freely accessible online from source government sites. This material should be as up-to-date and to the best possible quality and depth affordable. I also maintain that drafting offices should be the publishers and custodians of the statute book.

For better or worse I also straddle the historical and technological divide of paper and electronic publishing and the conflict that this often generates.

History and technology. I will try and provide a short history of the local online publication of legislation by source sites. I will portray this in the context of the ownership of documents and government policies. The subject will necessarily venture into some of the key technologies, which I will briefly describe.

Future developments and collaboration. I will also describe some of the likely directions for government legislative publishing and what the impact might be on the traditional paper products.

Similarly I propose a development framework and a more collaborative approach for the next wave of online developments.

2 Brief history of online access

2.1 Early history from 1990

Parliamentary Counsel Offices and Government Printing Offices. In the early 1990s the traditional government printing offices started to decline and even disappeared in some cases. Desktop publishing enabled many drafting offices to bring the pre-press function entirely in-house leaving just the bulk printing and distribution to external or contract printers.

Integration of functions. This development led to the growth of in-house publishing and IT skills in drafting offices and the closer integration of the drafting function with publishing. This stemmed from a logical desire to be able to produce the finished product as a single process rather than surrender control of the documents and then have to repeatedly review and proofread them. There was an increase in interest in simplifying the typography and production of printed legislation partly as a result.

Technical development. The key technical ingredient in the development of online publishing capacity was the capture of legislative data by drafting offices or in some cases printing offices. Most drafting offices started by using WordPerfect, later switching to Word. Quite sophisticated template systems for drafting and formatting legislative text were developed and exchanged across some jurisdictions.

2.2 Mid history▣

Internet. The growth and acceptance of the Internet in the mid 1990s and the enthusiasm by governments for online delivery of services coincided nicely for drafting offices. Most PCOs had a good start in terms of technological readiness. Also, clear government policies are a great help when preparing business cases for IT funding.

Copyright relaxed. In NSW, copyright in legislation was waived and this helped to demystify and de-commercialise electronic publishing. The PCO made its WordPerfect source data available for the first time to commercial publishers and AustLII.

AustLII has played a leading and catalytic role in gathering data from all of the jurisdictions. This has been a remarkable exercise in perseverance and diplomacy – to acquire and then render this material into its collection of databases and to continue to champion the cause of free online access.

E-government. Electronic publishing and the growth of the use of the Internet for research in the 1990s converged nicely with government interest in using online technology. All jurisdictions have had policies to encourage online services. These become infectious– perhaps driven by the interest in making more information freely available but more probably to speed up transactions, revenue collection and reduce costs as well as keeping in step with the latest technology elsewhere.

IT projects and resource switch. Significantly, PCOs began to switch the resources they had developed over many decades. The effort to maintain hardcopy mastersets of legislation (used for internal reference and reprinting purposes) was switched to capturing and maintaining electronic consolidations. The skill set for editorial production staff began to require above average IT skills. The projects in drafting offices based on more sophisticated technology gained a higher profile with capital costs of well over a million dollars.

2.3 Recent history▣

New publishing technology has been adopted unevenly in Australasian jurisdictions and elsewhere. Less than half of the drafting offices have or are moving to XML based systems (NSW, South Australia, Tasmania, New Zealand). Most jurisdictions are drafting and publishing using Microsoft Word and Adobe PDF.

Integration: websites and process. The convergence of technology and the legislative process is an interesting development and has not received much attention. The Australian Capital Territory has made its legislation register/website part of the legislative process. This movement - to authorise source databases and websites is increasing: the Commonwealth's FRLI and the proposed NSW e-gazettal process are examples. The collection and online publication of legislative metadata should also result in some other new products.

3 Ownership and policy underpinnings■

The ownership and control of documents are pivotal in shaping centralised legislative drafting and publishing. The more players and document transfer involved, the harder it is to build systems for effective online access. If a whiff of copyright or revenue is injected into the mix, it is likely to complicate matters further.

Role of a PCO. There is also still debate among some drafters about the role of a PCO. Perhaps some lingering desire for a simpler environment in the days when writing the words down was the only task and someone else worried about publishing them. In reality, most drafting offices have always had the role of performing some kind of consolidation or reprinting process. PCOs have kept tables and indexes tracking the legislation lifecycle and most importantly have had the special skills and knowledge to maintain the integrity and quality control of this work.

In almost all Australasian jurisdictions, the roles of drafting and publishing legislation now comfortably mesh with the maintenance of an electronic statute book and websites. There is still some discussion about how far this work should go. Some might claim that it strays into that rather imprecise zone known as “value-adding”. In the UK for example the historical versions of legislation are considered to be a value-added product. These will not be freely available from the new government online service.

Copyright and revenue. Copyright in legislation was waived for NSW in 1995. There has been free public access to online material for many years and it has raised very few copyright issues. The precise regimes vary from jurisdiction to jurisdiction (in some there is no copyright in legislation at all). It seems highly unlikely that public online access would become limited by this in any way in future. The revenue potential from source legislation in Australasia is low, given the relatively small size of the jurisdictions and because there is now so much free access.

E-Government and government policies. E-government has been a catch-all policy area for several years especially in terms of online access to government services and information. Payments and licensing are now probably the more popular areas but the provision of free public access to legislation has certainly been well lubricated by these policies.

The legislative process itself however remains a robustly paper driven in most jurisdictions. The passage of Bills and the making of statutory instruments still follow a nineteenth century document trail.

“E-assent” where the Governor applies an electronic signature to a document passed around on a secure network within government after parliamentarians have debated and even amended it using laptops in the chamber seems unlikely in most jurisdictions.

The authorisation of online versions is also likely to be a gradual development.

However, online legislation will inevitably become more integrated with the actual process as more practices such as the ACT and Commonwealth registers and NSW e-gazettal emerge.

4 Ownership & control examples

4.1 S. E. Asian illustration

Ownership and control of legislative material and its accessibility are not always so clear.

- Cambodia provides a florid example.■

I worked there on a project in 2004 to find ways of improving access to source legislation and judgments. There was no central drafting agency or co-ordinating body and the more powerful Ministries engaged their own drafters and publishers. Foreign influences were pervasive. ■

Ministries and some courts were loath to part with their documents, which had the potential for being sold separately and selectively, rather than being collected and distributed centrally. If legal publishing and distribution were performed at all it was a matter of expert local knowledge to find it. Paper documents with a suitable number of stamps and seals were copied and recopied until they were barely legible. These were sold to the cognoscenti both over and under the counter. The electronic version was generally not retained or archived and was often just deleted when the typist's hard drive became too cluttered.

The organisation producing the Government Gazette was poorly resourced and housed and relied upon a bank of ordinary office photocopiers.■ Two kilometres away was a state of the art, government owned printing factory operated by another Ministry that could have printed and distributed all legislation in a few hours each week.

Regrettably, Cambodia in 2004 was a place where all services and transactions seemed tightly controlled because of their private value potential.■

4.2 Local illustrations■

Fortunately for most of us, our part of the world is far less colourful and challenging. However there are many subtle differences in control that have shaped access:

- Western Australia still has a robust State Law Publisher that performs the traditional role of the Government Printer
- New Zealand outsourced its legislative database and legislative printing operations to the private sector and has had to reclaim these and build its in-house expertise from scratch. Its PCO is also part of the parliamentary campus.
- The ACT is a very small jurisdiction with a close-knit bureaucracy on a single secure network.
- Commonwealth. Has two drafting offices, one for Bills and the other for subordinate legislation and publishing. For Bills, the Parliament also controls the document during the passage stages.
- Victoria's Parliamentary Counsel also has the role of Government Printer.

These local differences have shaped the way that legislative publications are produced and managed as well as affecting the levels of accessibility. Also, the way a drafting office is established and fits within the local bureaucracy can affect its operations and budget - some are more independent than others in terms of the choice of technology and staffing arrangements.

5 Technology

5.1 Word processing to electronic publishing & XML

Legislative drafting offices are slowly moving towards more sophisticated systems for drafting and publishing. The size and complexity of legislative instruments test the boundaries of standard word processing applications. The dynamic nature of legislation and the frequency of amendments require long-term storage and tracking. The lifespan and structural complexity of the material are quite different to that produced using desktop publishing applications suited to one-off publications such as annual reports or prospectuses.

SGML & XML technologies (I'll lump these two together as XML or eXtensible Markup Language)

XML has much to offer:

- Portability, inter-usability, longevity of data
- Publishing in paper + multiple electronic formats + rich metadata from single data source
- Automated HTML provides fast internet accessibility + widest capacity to hyperlink and search
- Freedom from proprietary software?

5.2 Data capture and markup

In addition to the underlying technology, data capture and mark-up for legal drafting and publishing purposes are also complex and costly.

Analysis. The data has to first be fully analysed and a document type definition or schema prepared. This document identifies all the elements, numbering systems and references that are to be found in the data. This gets increasingly complex the greater the historical span of legacy data involved.

Capture. The data has to be captured by double keying or conversion if it is already in a reasonably consistent electronic format.

Markup and QA. The text is marked up to the document type definition. The specifications and standards have to be minutely described. QA systems need to be devised and run by the client to get the necessary level of accuracy. The current methodology is based on a standard: ISO 2859-1 and involves processing and testing manageable batches of material. This can often be an iterative process and it pays to start with batches of more straightforward material.

XML sample. The sample illustrates some of the potential of highly marked-up XML data. Every heading and element are identified and tagged. This is also within a closely defined hierarchy and not just styled as in word processing documents. The cross-references to other legislation are marked up, as are the individual penalties and where they are nested. This enables a high degree of automation to occur when

reusing the data internally, publishing the data online in HTML or when moving to other applications. For example, the NSW Judicial Commission takes the PCO's XML data and links this with its sentencing databases, relevant cases and bench books. The material is then published on the Commission's Intranet, using open-source XML applications. One of those rare examples of highly effective inter-agency collaboration. The same data is provided to AustLII, which finds the format and quality superior to most data it republishes.

Offshore expertise. ■ Much of this work is now done off-shore. India has a lot of expertise in this field and the legislative data from NSW and South Australia was marked up by a company in Mumbai that also does similar work for commercial legal publishers.

5.3 Metadata ■

Metadata is the information about the document and for legislation it is especially critical for electronic storage and publishing. The minimum needed for identification purposes is:

- Title (and any name change), year and number
- Classification (ie instrument type)
- Assent and commencement dates
- Currency: ie in force or repealed etc, date of repeal/repealing instrument

The potential in terms of the extent and reuse of metadata is far greater:

NSW PCO collects up to 50 items of information and status events about each piece of legislation. These are captured as metadata within its XML-based systems and about half of these are public events that could be provided online in some way. For example the dates of notice of motion for a Bill, introduction and passage dates, details of amendments in committee, assent and proclamation details, any review dates, the subsequent amendment history and repeal. This could provide users with more information about the passage and lifecycle of legislation. More significantly, it should be provided in more dynamic ways than just serving up the usual static tables and indexes that have to be located and consulted as secondary sources.

5.4 XML systems: requirements and overheads ■

Full XML legislative systems are not for the faint hearted. The quality control and documentation in building these systems are formidable. There are very few IT experts that have the necessary specialist knowledge of both XML and legislation. The standard approach of throwing IT resources at the task doesn't seem to work and development takes a long time. It requires in-house expertise working along side the external specialists. The costs and risks are quite high - these are often multi-million dollar projects - although the benefits in internal operations and public access are clear enough. These will be more obvious in future years when the depth of point-in-time collections becomes more extensive and there are more options for the way legislative data can be presented and linked. These systems would be hard to justify at present for entirely static, large archival collections.

5.5 Scanning: accuracy and quality control, NSW experience

Drafting offices tend to be mesmerised by accuracy and detail so any form of data capture is anxiety inducing. Even full proof-reading of documents is sometimes considered inadequate and this is a cautionary tale about scanning.

NSW scanned a decade of annual volumes of both statutes and regulations about five years ago. These have been online on the legislation website for some time although the scars are still raw. Volumes were guillotined and scanned in using the programme Adobe Capture. This captured the actual characters so the documents would be searchable and not too large in size. I have to point out that it was an early version of the software and the technology has moved on considerably. I plan to revisit this area in the coming year.

The initial capture part of the exercise was straightforward but the quality control was extremely laborious, inspecting each page of text even though the material was from the 1990s and fairly clean. The most significant feature about the format of legislation is that presentation and indentations impart meaning and its capture is not as simple as capturing works of fiction or standard text books a la Google.

■ **General problems** encountered were with:

- text order, paragraph blocks, hanging indents
- lines out of order
- originals too dark or light and creases
- forms, tables and images (these are tricky in any application)

Then there were the **character recognition** problems:

- ligatures and similar: ff, ll, ur, rn “Goulbum”
- italics
- distinguishing between bold and Roman

Manual inspection of each page and the numbering sequence requires intervention to correct text, which introduces another element of risk.

Sample■ shows a scanned amending rule where the paragraph blocking has been misread and the two, lettered paragraphs have appeared at the end.

5.6 Scanning older material■

Scanning 19th century material is more challenging and needs careful inspection before scanning:

- Frail paper, rust spots, marks and even ink annotations
- Damaged pages or type
- Multiple typefaces, extra large dropped capitals, caps and small caps, bold italic
- 8 point sidenotes, some over 200 characters long
- multi-column, multi-page tables including landscape tables
- variations in print design, ie contents tables

Large-scale capture projects of historical legislative material throw up issues of ownership and management – should it be libraries, archive authorities or drafting offices? Also how to integrate these collections with current material?

5.7 Future capture

It is unlikely that full data capture and mark-up of large quantities of historical statutes are affordable at present. There is a tipping point when large quantities of what is often fairly static data become too complex or expensive for the ideal capture solution. Cheaper scanning processes at least can provide a form of wider access – half a loaf is better than none. However even scanning statutes still needs detailed analysis and QA as well as some manual metadata entry to make it reasonably worthwhile and accurate. If this is done well, there is always the possibility that the data can be developed in future.

There is growing interest in capturing archival material across many fields and jurisdictions. NSW is very interested in pursuing scanning technology and in possible collaboration with others.

6 Paper products

6.1 Static, unchanging range

Official legislative publications generally develop very slowly and haven't changed for years. The usual suite comprises:

- new legislation in pamphlet and bound forms, sometimes also a Gazette
- point in time consolidations or reprints of principal Acts and regulations
- information guides listing what legislation has been made or amended over a period – week, month or year
- information guides listing what is in force or repealed at a point in time.

6.2 Limitations, numbers and costs

There are now less than 250 subscribers to printed new NSW Acts and similar numbers for Reprints. There are less than 100 for regulations and various information guides. A set of bound NSW statutes and regulations requires about 30 linear metres of shelving, the Gazette over 70 metres.

The full suite of products must cost in excess of \$6,000 a year to maintain and of course remain static in nature and traditionally slow to print and distribute. These printing, retailing and distribution processes are increasingly likely to be out of the hands of traditional GPOs. PCOs do not usually control these activities.

6.3 Future?

While there are established libraries and the legislative process itself remains paper driven, the current paper products are likely to be maintained. NSW plans to maintain its production of both paper and online services for the time being, but ultimately it does not control the printing, sale and distribution of hardcopy legislation.

As subscribers continue to decline, what will the paper costs and production issues be when there are only 100 or so left?

The recent demise of the annual bound Commonwealth SR series is noted.

7 Online products

7.1 Infancy

In contrast, many online legislative publications and particularly the source sites are relatively young and clearly still evolving.

The NSW site www.legislation.nsw.gov.au was launched in June 2002. The South Australian site is expected to be launched any day now.

Apart from the constant maintenance of the data on the NSW site, historical material is being back captured and added. Some new features have been added already and many more are planned. We want to be able to publish applied laws, deemed statutory instruments and readable maps.

7.2 New products

Looking at new online products, the **Weekly Bulletin** is hardly new, it's almost 7 years old and has built up over 2,000 subscribers. For those not familiar with the bulletin, it's a very simple summary of all legislative events for the week and purely an online product. Significantly, it has the scope to be linked with all of the new instruments that have been made. This is possibly the model for the proposed e-gazettal. It would enable the subscriber to have immediate access to all new legislation – delivered to the desktop – instead of having to track and download individual instruments. A “notify me when” service is a similar concept that we are looking at.

“Catalogue” This is a list of all New South Wales Acts reaching back to 1824. It will initially contain key metadata about each title but will have the capacity to add the full text when the resources and technology are found for this task. The titles and metadata have already been captured and we are investigating ways of presenting this to users and integrating it with the current website. This is an area where we would particularly welcome input from librarians and specialist researchers.

7.3 Integration with legislative process

Earlier this year the NSW Interpretation Act was amended to provide a statutory basis for the legislation website and to provide for the on-line gazettal of statutory instruments. This will include a wider range of instruments than currently appear on the site. The amendments also enable the Parliamentary Counsel to certify the form of on-line legislation that is correct and provide it with the same status as the current paper reprints. These amendments are yet to be fully proclaimed as they are linked to the completion of a document management and tracking system and some changes that need to be made to the website.

This type of development, where the actual legislative process becomes more closely integrated with the source website, is likely to be adopted elsewhere.

The ACT already has a legislation lodgement system that is part of the legislative process, underpinned by its 2001 Legislation Act, and it certifies the online versions.

The Commonwealth has the Federal Register of Legislative Instruments (FRLI) and provides authorised, source versions.

New Zealand and South Australia have plans to authorise online versions at some stage in the future.

These online developments will inevitably affect the traditional hardcopy publications.

Apart from cost implications, there are the “look and feel” issues associated with the online product. Traditionalists object to the loss of the B5 size format and the richer and predictable typography of printed legislation. There is more work needed to improve the online products in this respect.

7.4 Development path

Non-disruptive. A major challenge with relatively complex websites is how to add more features but not disrupt regular users or reduce functionality (or content!)

Incremental back capture of data and more linking are enhancements that will not dramatically change the look and feel of the site. These enhancements will include more historical statutes and point-in-time versions, possible linking to Hansard and other related documents. Similarly, a notify me when facility is not difficult to integrate (although it poses some operational challenges in terms of whether it notifies actual legislative events as they happen or merely when the status of a document on the system changes).

There needs to be more and better use of metadata, for example the historical clause by clause tables that currently appear as text at the end of the Act or instrument should be in a linked and more reusable database (as already occurs on the Tasmanian site).

The addition of features such as e-Gazettal and a more user-directed catalogue for all titles past and present are quite challenging to integrate with the existing site, apart from the effort required to capture historical data and metadata.

A development framework of 5 to 10 years is probably required to complete all of this.

8 Collaboration

8.1 Record so far

The history of collaboration in this field has not been impressive. The various drafting offices do meet at least annually and there is the Parliamentary Counsel's Committee IT Forum that has existed since 1995. This enables a good deal of technical information at the operational level to be shared. There have been similar approaches taken by several jurisdictions in the redesign and formatting of hardcopy legislation but most have gone their own ways with websites and on-line delivery. The DTDs developed for NSW have been adopted by South Australia and are freely available for others to use. I sometimes speculate what could have been done if we had all pooled some resources 7 or 8 years ago and developed a common system and website?

8.2 Law librarians

The development of new online features will need the involvement of law librarians as the key users - perhaps the only other group that really understands the finer details of how legislative publications work.

NSW will be inviting you to participate in testing and to familiarisation seminars as these features are developed and released.

I hope that there will be many more opportunities for both groups to attend each other's conferences and collaborate on some of these new projects.

Thank you for this opportunity and please feel free to contact me, I welcome questions.