01.

Advanced tools for legislation

Timothy Arnold-Moore

01 Advanced tools for legislation

Dr. Timothy Arnold-Moore[*]

Exploring the possibilities and limitations of advanced tools for editing and managing legislation.

I Introduction

Collections of legislation contain some of the most important documents in any country. The nature of legislative drafting processes and the documents they result in create requirements for unique and specific sets of tools for managing and publishing them. The characteristics of the drafting and enactment processes, the documents produced by them, the financial and intangible benefits of improving the drafting and publishing processes, and the availability of legislative material justify a significant investment in such purpose-built tools.

The fact that the legal community benefits most directly from delivering better publishing outcomes does not undermine the broader community benefit of such an investment:

> "It cannot be doubted that dissemination by publication of accurate copies of statutory enactments is beneficial to the community as a whole; and this is not the less so because at least in many

* Dr. Arnold-Moore holds a Bachelor of Laws from the Melbourne Law School and a Bachelor of Science (Hons) in Computer Science from the University of Melbourne. His Doctoral research and thesis "Information systems for legislation" from the Royal Melbourne Institute of Technology lead to a career of over 20 years designing and building information systems to support drafting and publishing legislation for governments including Tasmania (EnAct), New South Wales (LEGIS), Singapore (LEAP), Canada (LIMS) and Queensland (QuILLS). He has also consulted to drafting offices in New Zealand, Ontario and the US House of Representatives and presented to conferences of legal professionals in Washington, Hobart, Sydney, Chicago, Ottawa, Rome, Tokyo, Wellington, Kyoto, Melbourne, the Hague, Brussels, Boston, Cardiff, Canberra and Seoul. Dr. Arnold-Moore now works for Leidos Pty Ltd, the Australian subsidiary of global technology company, Leidos, Inc.

> instances the ordinary member of the public either does not attempt to, or cannot by study, arrive at a true conclusion of their import, or because the true understanding is largely limited to persons engaged professionally, or as public servants in the field of any particular enactment or otherwise interested in that field ... To state that the publication also provides many professional men with the tools of their trade does not seem to me in any way to detract from the benefit that accrues to the community from the fact that the law does not remain locked in the bosom of the Judiciary." per Russell LJ *Incorporated Council of Law Reporting for England and Wales v. A-G* [1992] Ch 73; [1971] 3 All ER 1029, 1034.

This paper seeks to demonstrate the desirability of such tools, describes a number of such tools, and the benefits that they can provide to both drafting offices and consumers of legislative material. As a commercial provider of such tools, I must declare something of a conflict of interest here. However, the views in this paper borrow widely from others within the legislative drafting community, broader government policy positions, and legal academia.

Ⅱ Objectives for IT systems supporting drafting and publishing legislation

Why legislation is different

A legislation drafting and management system on the face of it looks like a typical office document authoring, management, and delivery application. But there are a number of factors related to legislation that distinguish it from typical office environments. These include:

- **Time factors:**
  - **The longevity of the documents:** Some legislation being managed today was drafted hundreds of years ago – portions of the Magna *Carta* (1215) are still part of the law in many English-speaking countries. By contrast, most office documents are created for immediate use. Few are retained for more than a year or two.
  - **Frequently amended:** The vast majority of legislation made is amending legislation –

that is, the wording of the legislation describes textual changes to existing legislation. While these amendments have the force of law, users of legislation care more about the wording that results from applying the amendments to the existing substantive provisions (referred to as "consolidating" the amendments, hence the term "consolidation" to describe these documents) than the actual text describing the amendment. As a result, there are often multiple versions of the substantive legislation with a series of versions each valid at a different time. Users of legislation require immediate access to the most current consolidation of a piece of legislation, but they also need access to past versions – when preparing legal advice or hearing a case about a past incident – and, to the extent possible, future versions – when preparing advice about future activity. Office documents rarely exist long enough for anything but the latest version to be of interest. The time period of validity of most office documents is rarely as clearly defined as for legislation.

- **Nature of documents:**
  - **Consistent naming and formatting**: Legislation in most English-speaking jurisdictions, follows largely the English tradition breaking Acts into numbered sections, subsections, paragraphs, and subparagraphs each with distinctive numbering conventions, typographic markers, and consistent nomenclature (known as "paragraphing"), and grouping the sections into chapters, parts, subparts, divisions, subdivisions, and under various unnumbered or unnamed headings. While some of the typographic markers have changed subtly over the years and vary from jurisdiction to jurisdiction, the structure is largely unchanged and remarkably uniform across most English-speaking jurisdictions. Civil law jurisdictions have a little more variability, but the concept of Articles broken up into sub-articles, clauses, and paragraphs, or grouped in parts and chapters is common to most civil law jurisdictions. Office documents vary much more widely in structure both within a corpus and between corpora.
  - **The importance of the documents**: The documents recording the legislation of a country are amongst the most important documents in the land. This importance justifies

particular care in ensuring the highest quality of print production when rendering the documents. The main driver for office documents is efficiency. Providing that the documents look sufficiently professional, minor variations in typography can be overlooked. While efficiency and timeliness are important for legislation, presentation is particularly important and variations in typography, however small, may be noticed and trigger unnecessary delays in the legislative process. This leads to a justified culture in drafting offices of exacting standards and a focus on minute detail.

- **Time pressures placed on the drafting offices**: While the predictability of a legislature varies from jurisdiction to jurisdiction, and parliament to parliament (sometimes depending on election results within a jurisdiction), most drafting offices have periods when they are working on an urgent draft right up until tabling in the legislature. Committees may work on amendments and these amendments need to be consolidated so that the Bill for third reading can be presented very quickly. The results of amending legislation need to be available as soon as possible after enactment, making, or commencement. Therefore, drafting offices are under considerable pressure to prepare paper and electronic versions of a Bill or Act very quickly.

- **Nature of the creation process:**
  - **Highly regulated enactment process**: While the exact process of enactment varies between monarchies and republics, bicameral or unicameral legislatures, legislative traditions, and between regulation and legislation, the common thread is that legislation only comes about if the required procedures for making it are followed. Until then, it is merely a wish of its sponsor. Many drafting offices also have a formal internal sign-off process to ensure quality before documents are sent out of the office to the legislature or government departments.
  - **More fluid drafting process**: By contrast, the initiation of a drafting project is often quite informal. In Westminster governments, government Bills often require Cabinet approval to begin drafting or to table a Bill. Typically, a request for a draft regulation is less formal. Even in jurisdictions requiring Cabinet approval to draft, the process may be

initiated in anticipation of receiving that approval. Instructions may be received once or continually during the drafting process with drafts flowing out, instructions (and possibly proposed drafts) flowing in. Different drafters may work on a project at different times or be called to review some or all of a draft. Significant flexibility is required during the stages leading up to tabling for enactment.

- **The unique relationship between the drafters and the legislature**: The drafters provide legal advice to the legislature and owe fiduciary duties, including a duty of confidence, to their clients – the sitting members, the Ministers, and their staff. The documents, while being drafted, are highly confidential and may contain politically sensitive (and occasionally militarily sensitive) information. Once the drafts are tabled, they are public documents. In a normal office environment, most documents either stay sensitive throughout their life cycle or are designed from the beginning for public distribution.
- **Split functions of drafting offices**: Drafting offices such as Parliamentary Counsel or Legislative Counsel have existed for many years solely to draft bills and associated material and regulations and other subordinate legislation for the government. However, there is an increasing and logical trend for such offices to take over the role historically taken by a government printing office of publishing legislative material. Such offices are uniquely positioned to maintain many of the products useful to the broader public for its own internal use (including up-to-date consolidations required for preparing amendment Bills or draft Regulations). A culture of providing a discrete and largely confidential service to the Government does not always sit well with the more public service of publishing.

Resulting internal considerations for legislative tools

Since the structure of documents is regular and both the documents and the structure are long-lasting, a significant investment in tools to create and manage legislation is more easily justified than with most office documents. This investment can take the form of:

- purchasing specialized licensed software,

- customizing that software for particular tasks,
- migrating from one format to another (and the resulting checking of that data required), and
- the investment in training staff to use the new tools (with the inevitable dip in productivity immediately after such tools are introduced).

If an office drafts 100 Bills and 1000 subordinate instruments in a year, a tool that can save 1 hour in the preparation time of each document has a fairly simple cost-benefit analysis. Saving an hour of the legislature's time saves the cost of the legislature salaries and those of all their support staff and infrastructure along with the lost opportunity cost of what might have been achieved with that hour on some other useful project. Saving a week or a day or an hour in making a consolidated Act available to the public is much harder to measure in financial terms, but ready access to timely legislative information can significantly reduce the costs and improve the timeliness of the provision of legal services. I'll leave a formal analysis to the economists, but the cumulative cost of maintaining up-to-date paper libraries of legislation in government departments, law libraries, and legal offices in a small jurisdiction such as Tasmania was conservatively estimated at well over $5 per person living in the jurisdiction per year. Clearly, a significant investment in high-quality tools to support the legislative process and deliver legislation and related materials to the widest possible audience in a timely manner is justified.

With a large number of drafting projects typically managed by a drafting office with complex stages and strict quality assurance and legislative processes, it becomes necessary not only to perform all of the required steps, but also to prove after the fact that those steps were followed. The value of this governance in financial terms is hard to quantify, but an efficient and effective workflow management system and archiving process can reduce the cost of compliance to a drafting office significantly as well as improve their standing with their constituents.

Tracking each and every process also facilitates better management practices within an office as management can more easily monitor the work-load on staff and better distribute and prioritize tasks to ensure the most efficient delivery of their services. In jurisdictions where individual members consume government drafting resources or government departments are

required to account for the burden they impose on the drafting resources, such tracking allows management to 1) quickly determine (and manage) overly demanding clients, 2) account for costs, and 3) justify changes in resourcing.

Resulting external considerations for legislative tools

External considerations for legislative tools are typically grouped around either:

- **Governance**: ensuring the right steps are followed by the right people and proving that they have been after the fact; or
- **End user demands**: availability and suitability of the legislation for different uses and purposes.

The value of governance and business process tracking or workflow management extends well beyond the office also. The legislature and the legal system also need confidence that the legislative materials they are dealing with are correct and up-to-date. The ability to demonstrate that correct processes were followed to create legislation, apply amendments, and make them available is of significant value beyond the drafting office.

End users often request (or demand) additional formats, information, or metadata related to legislative documents. This often requires a change in the underlying source format of the legislative documents. The importance of the documents creates an impediment to changing the format of the documents. Every time the source format of the document collection is changed, the results have to be checked. The cost of this checking depends on the degree of change. Migrating from one version of Microsoft Word to another may require little checking of the document contents, but may involve significant changes to macros and other customizations. Migrating from paper or pre-existing electronic formats to XML requires significantly more effort to migrate, but has the advantage that further data migration is unlikely to be required in the foreseeable future. A proprietary format, such as a native word processor format, requires update when there are changes to the software version, software component, or house style for formatting the legislation. By contrast, a structured XML format

is independent of the software version or software component used for the relevant task, be it authoring or formatting the legislative documents. Changes in formatting across the entire collection can be managed largely with changes in the stylesheet sets (that define the formatting of particular structures within the legislation collection). Such a migration also "future proofs" the data collection as new stylesheets can be added to generate the next desired format whether it be the latest version of a popular word processing package, the latest new enhanced HTML, or formats for new devices such as eBooks which may not have existed when the data was originally migrated.

Potentially onerous tasks for drafting offices

External consumers of legislation often wonder why more isn't done with these high-value documents by the legislative drafting offices. While I have written about constitutional reasons for avoiding some categories of "value adding" by particular government offices elsewhere[1], I suspect the reasons are more mundane and obvious.

Data migration is hard and expensive in any context.[2] In addition, because of the importance of the documents, it is also more resource intensive in a legislative environment. The converted results of legislative documents require extensive checking against the original or, input, documents. This review process is required when systems are updated or prior to publication.

Subject classification of legislative documents is extremely useful to consumers of legislation, particularly those with limited familiarity with the corpus. While it supports a conventional "browse by subject", it can also be used to support faceted browsing of search results[3] and to create usable subsets of legislative materials for use by particular government departments or user communities.

Cross-references to other provisions within a legislative document, other documents, or

1) Arnold-Moore, T "Public access to legislation and the democratic process" [2004] 5 *RegelMaat* 162.
2) Ebrahim, Z & Irani, Z "E-government adoption: architecture and barriers." (2005) 11 *Business Process Management Journal* 589.
3) See Arentz, W, *Searching and classifying non-textual information* 2004 Doctoral Thesis, Norwegian University of Science and Technolog; Arentz, W "Multidimensional visualization and navigation in search results" (2004) 3213 LNCS 620.

provisions within the legislative corpus are common.[4] Ideally, these links will be marked appropriately in source documents so that web navigation of the legislation collection supports hypertext link behavior corresponding to the cross-reference wording.

Government is eager to be seen as reducing the regulatory burden on business and citizens.[5] This leads to a need for governments to monitor progress in increasing the efficiency of or reducing regulation. One way of reporting this progress is to simply count pages of regulation by department or agency and score them on reducing the page count. Other, more sophisticated, methods count regulatory complexity. This task is often pushed to the legislative drafting office as they manage a complete collection of the regulations and have the most immediate knowledge of changes in the collection. Regardless of the method used to measure regulatory complexity, some effort is required to classify which regulations are managed by which agency and to compile the measures across the regulations.

To summarize how these tasks are onerous on a drafting office:

- **Creating numerous formats at time of publication:** While it is desirable to make a wide variety of formats available for government and public consumption (including HTML for easy browsing on electronic tools, PDF for faithful reproduction of authorized paper publications, XML for longevity and future formats, and eBooks for tablet and smartphone delivery), labor intensive processes to create them at crucial times are difficult to justify.
- **Adding subject metadata to documents:** Assigning a controlled or open vocabulary of subjects to legislative documents (or provisions within them) is a complex manual task

4) See Arnold-Moore, T *Information systems for legislation* 1998 Doctoral Thesis, Royal Melbourne Institute of Technology, s. 2.2 citing Corbett, M "Indexing and searching statutory text" (1992) 84 *Law Library Journal* 759; Hoey, M "The discourse properties of a statute" in Walter, C *Computer power and legal language* (1988), Quorum Books: New York; Turtle, H "Text retrieval in the legal world" (1995) 3(1) *Artificial Intelligence and Law* 5.

5) See Gore, A *From red tape to results: creating a government that works better and costs less* (1993) National Performance Review: Washington; Keyworth, T "Measuring and managing the costs of red tape: A review of recent policy developments" (2006) 22 *Oxford Review of Economic Policy* 260; Lodge, M & Wegrich K "Crowdsourcing and regulatory reviews: A new way of challenging red tape in British government?" (2014) *Regulation & Governance* ⟨http://onlinelibrary.wiley.com/doi/10.1111/rego.12048⟩.

typically performed most consistently by highly trained law librarians. To require this analysis to be performed manually by drafting office staff, in addition to all of the other tasks that they need to perform in a timely fashion, at the time of tabling or publication is likely to slow the publishing process down or lead to low quality classification of the documents.[6)]

- **Encoding cross-references:** In order to support hypertext linking of cross-references in the text, the source format of the legislation (whether it is XML, Microsoft Word, or something else) needs to capture hidden information about the cross-reference such as the identifier of the referenced document and, if a reference to a provision within that document, the identifier of the provision. Managing and updating this markup in a document through the drafting process can be difficult in the face of the renumbering that often occurs when a heavily amended Bill is made into an Act, so this is often left to be accomplished at the time of publishing (if performed at all). Doing this manually can be incredibly resource intensive, again, typically at the time of publication when time is least available.
- **Tracking regulatory complexity:** As documents are added to the collection or amendments applied, legislative drafting offices are often asked to maintain statistics on regulatory complexity or page counts by responsible department or agency. These statistics are onerous to create from scratch but can be incrementally updated as changes are made and published. This creates yet another task at the time of publication for the drafting office.

This is a list of a few of the requested services that I am aware of. Despite the demands to do more, it is hard to require already stretched drafting offices, with significant time pressures, to perform extra tasks particularly at time-critical stages during drafting. The point at which many tasks are requested to be performed on legislation tends to coincide with the time that the drafting office or publishing team are under the most pressure to get material out to the legislature, a printer, or a website. These are the points in the process when time is least

6) Although I don't describe the process here in further detail, the NIR project has designed tools to address this problem – see for example Biagioli, C *et al.* "Automatic semantics extraction in law documents" in *Proceedings of the 10th International Conference on Artificial Intelligence and Law*. ACM, 2005.

available and when quality assurance processes are most important. Drafting offices are understandably reluctant to add any complexity or additional tasks at the time of tabling or publication that don't directly reduce the risk of error or improve the confidence in the correctness of the documents.

While all of these tasks can be incorporated into the legislative drafting and publishing environment effectively, careful choice of source format, drafting tools and processes, and targeted investment in customization are needed to make any or all of them feasible without a significant investment in human resources.

Any such additional work imposed for purely publishing purposes needs one or more of the following:

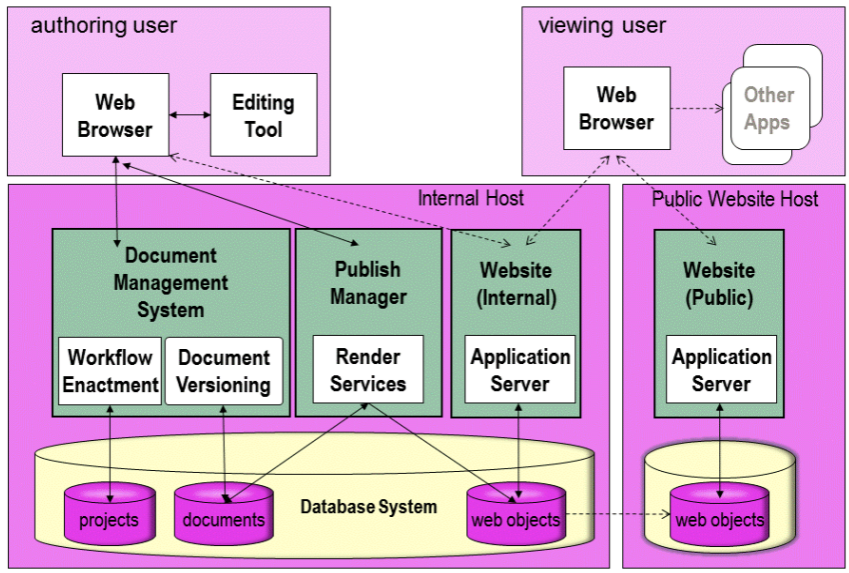
- a champion in the office who sees the value and can convince other staff members that the extra effort is justified;
- a tangible benefit for the drafting office staff – such as time saving for the staff or the office elsewhere in the process; or
- a top-down direction with the appropriate commitment of extra required resources.

The demands, features, and considerations noted above add to the complexity of constructing tools to support legislation drafting and publishing but also justify a considerable investment in such tools. These tools require both unique features and novel development techniques to ensure a lasting and usable system and to meet the demanding requirements of a legislative drafting environment.

Ⅲ Typical components of an IT system for legislation

Let us consider what components typically make up a system for managing legislative drafting and publication. For this purpose, I am largely ignoring the infrastructure around paper publishing of legislative materials for two reasons. Firstly, any jurisdiction considering a large information technology (IT) project to support legislative drafting and publishing either already has an existing infrastructure for paper publication that can be provided with source documents simply by delivering PDF documents to the existing infrastructure (possibly reducing or re-deploying any existing infrastructure required to prepare the formats of these documents). Secondly, an effective electronic publishing capability can significantly reduce the demand for traditional printed resources[7] as users can print on demand the versions of legislation they need in paper at a particular time and have all the benefits of interactive browsing and searching of the whole collection for minimal communication costs when a paper copy is not needed. A complete solution needs to consider the print production requirements and manage any change around these processes appropriately.

7) The recent scheme to deploy iPads (see *The Telegraph* 2 Apr 2013 "90 MPs get free iPads for working on 'paperless select committees'"; *The Independent* 27 Aug 2014 "MPs each get through a computer a year. Now they want iPads"; See also "Select Committees becoming paperless offices (2013)" 〈http://www.parliament.uk/site-information/foi/foi-and-eir/commons-foi-disclosures/other-house-matters/ select-committees-becoming-paperless-offices/〉) has reportedly resulted in huge savings in print costs.



[Figure 1] Typical architecture for legislative drafting and publishing system

Figure 1 above shows the architecture of a typical system to support legislative drafting and publication. Each of these components is described in more detail as follows:

- **Editing Tool:** A tool used by drafters and drafting office support staff to create new draft legislation and update existing consolidated legislation (reprints). In an XML environment, this is typically a structured editor such as PTC Arbortext®[i] Editor (used in New Zealand, Canada, and some U.S. states), JustSystems XMetal®[ii] (used in South Australia, U.S. House, Canada, and some U.S. states), Adobe® FrameMaker®[iii] (used in New South Wales, Queensland), or <oXygen/>®[iv] (used in some U.S. states). Some jurisdictions use a word-processor based structured editor (some U.S. states), a word-processor with automated translation (Tasmania, Papua New Guinea, Singapore [Microsoft Word], or Ireland [OpenOffice®[v]]), or manual translation (not necessarily within the office). Far more common is a purely word processor authoring environment – typically Microsoft Word (as used in Ontario, Victoria, Australian Federal).
- **Document Management System (DMS):** A DMS includes the following components:
  - **Workflow Enactment:** This infrastructure takes formal definitions of work processes

(drafting projects) and presents tasks to drafters and other staff based on the status of each instance or project. Other capabilities typically include reporting (static) and monitoring (dynamic) functions, alerts and escalation (to ensure deadlines are met), and records management/archiving.

- ○ **Document Versioning:** The legislative drafting process is a set of business processes designed to deliver documents. In the process, many versions of the documents are produced and refined before delivering the final version. Once enacted, documents are amended, so a series of versions of in force legislation also needs to be managed.

Typically, a DMS is highly customized with configuration for specific workflow definitions and management of particular versions needed at specific stages in the legislative process (e.g., first reading version, third reading version, copy for Assent, etc.). Typical software used in this environment includes PTC Arbortext Adapter® to Documentum®[vi], TeraText®[vii] DMS, or Microsoft SharePoint®[viii].

- **Application Server (Website):** Invariably, a large IT project to support legislative drafting and publishing includes either an internal or publicly accessible website or both. This requires an application server to support HTTP requests, provide browse and search logic, and access the various components and different versions of legislation provided by the website logic. This is typically Microsoft IIS or a Java®-based platform such as JBoss®[ix], WebLogic®[x] or WebSphere®[xi] although TeraText provides an application server used widely for legislative websites.
- **Database System:** Most document management systems rely on an underlying database management system either a traditional database system such as Oracle®, DB2®, or SQL Server®, or a NoSQL database system such as TeraText or SOLR™. In some cases, the file system provides the database system possibly augmented by a **Search Engine** to support text searching. Different components may use different database systems. The **Workflow Enactment** requires a database to store information about current and completed projects. The **Document Versioning** needs to store document metadata including version information and relationships between versions as well as the document content (sometimes stored

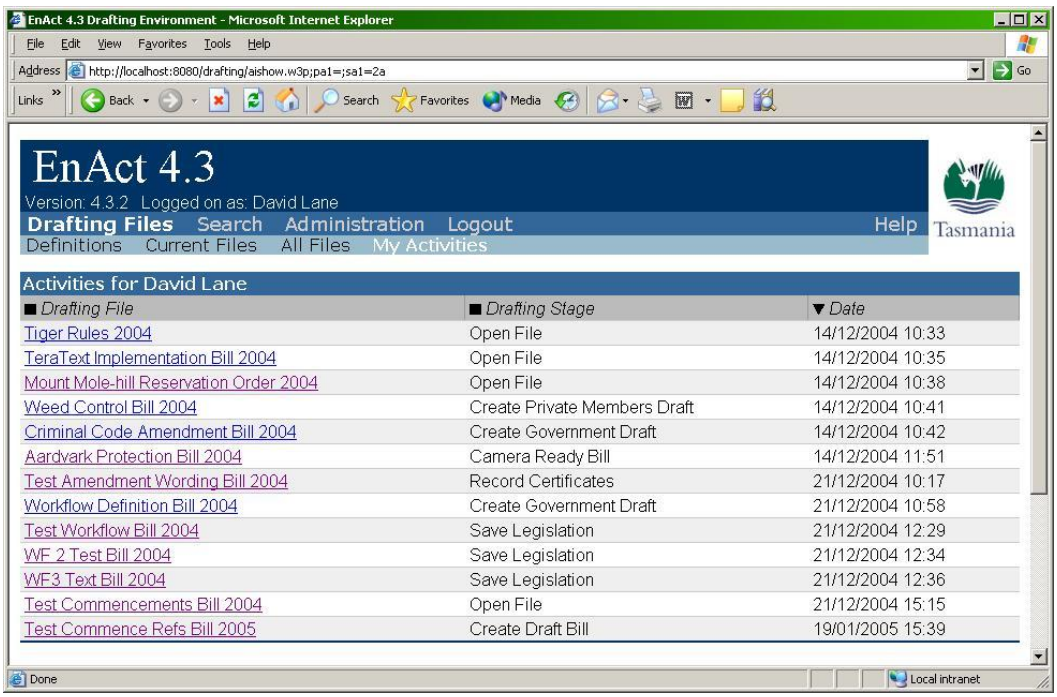
separately on the file system). The **Website Application Server** typically needs a repository of web objects to be delivered to the public – this may access the document version repository directly although this is more commonly a separate repository optimized for web delivery and may also be file system-based.

- **Render Services:** Certainly in an XML authoring environment but also in most other environments, it is necessary to provide a service to take the source format of the legislation to produce whatever other additional formats are required. In the case of XML, PDF rendering is important for producing print versions (if only for the legislative process), and HTML and other formats are important for web delivery including eBook and word-processor formats. For XML to PDF, commercial products such as PTC Arbortext Advanced Print Published (APP), Antenna House, Inc.'s Formatter, or RenderX, Inc.'s XEP Engine are used. In environments where a word processor format is the authoritative source, either the word processor itself or third party tools may be used to generate PDF, HTML, and eBook versions. In the former case, the Editing Tool fills a dual role. Adobe FrameMaker is designed to perform both functions for XML source material.

The indispensable components are the **Editing Tool** and the website **Application Server.** Low technology replacements for the other components are available but may involve reduced capability.

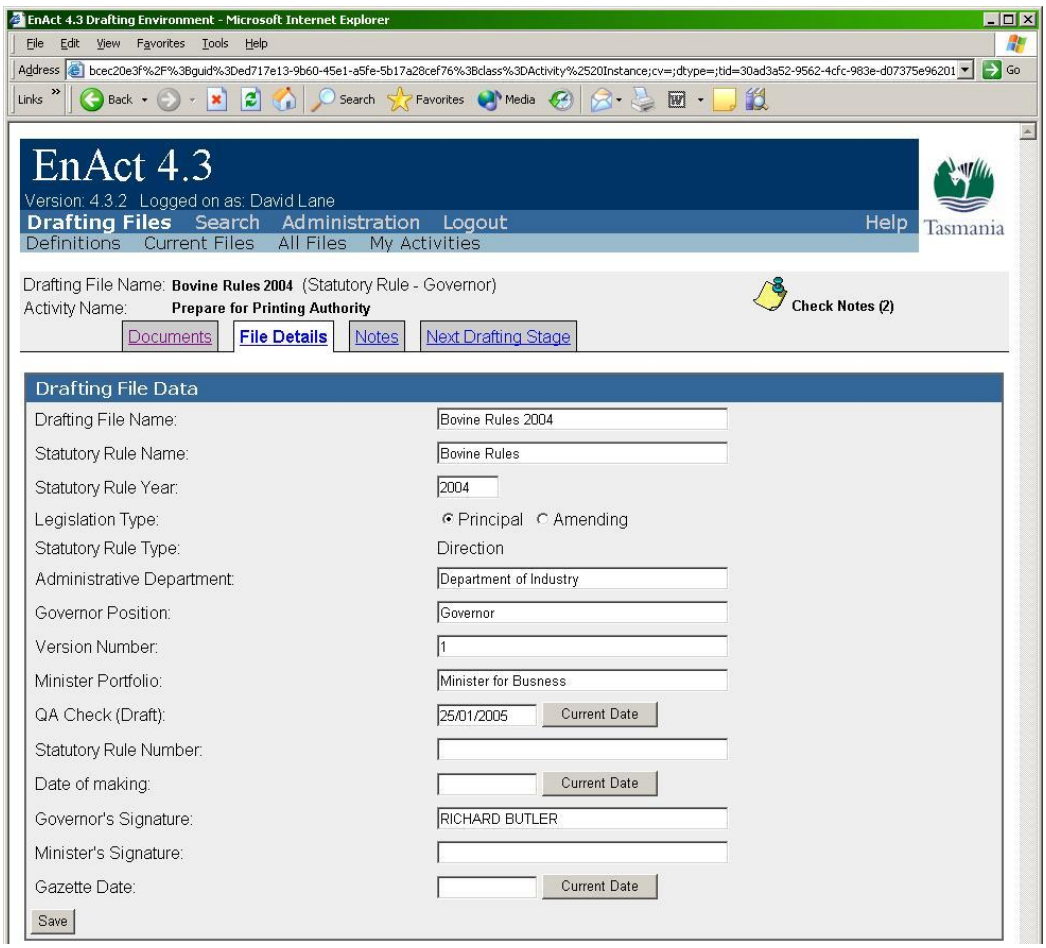
Some example pages from the Tasmanian EnAct Legislation document management system appear below. I have deliberately chosen to use earlier versions of the EnAct system to demonstrate that this functionality has been available for some time. Similar functionality using more sophisticated user interfaces, more modern web technologies, and more recent versions of software are available.

Figure 2 shows a list of Bill projects allocated to the user currently logged into the EnAct system. Note the title of the draft piece of legislation, the stage at which the drafting has progressed, and the date and time of the last change to the task. The list is sorted in ascending order of the last modify date.



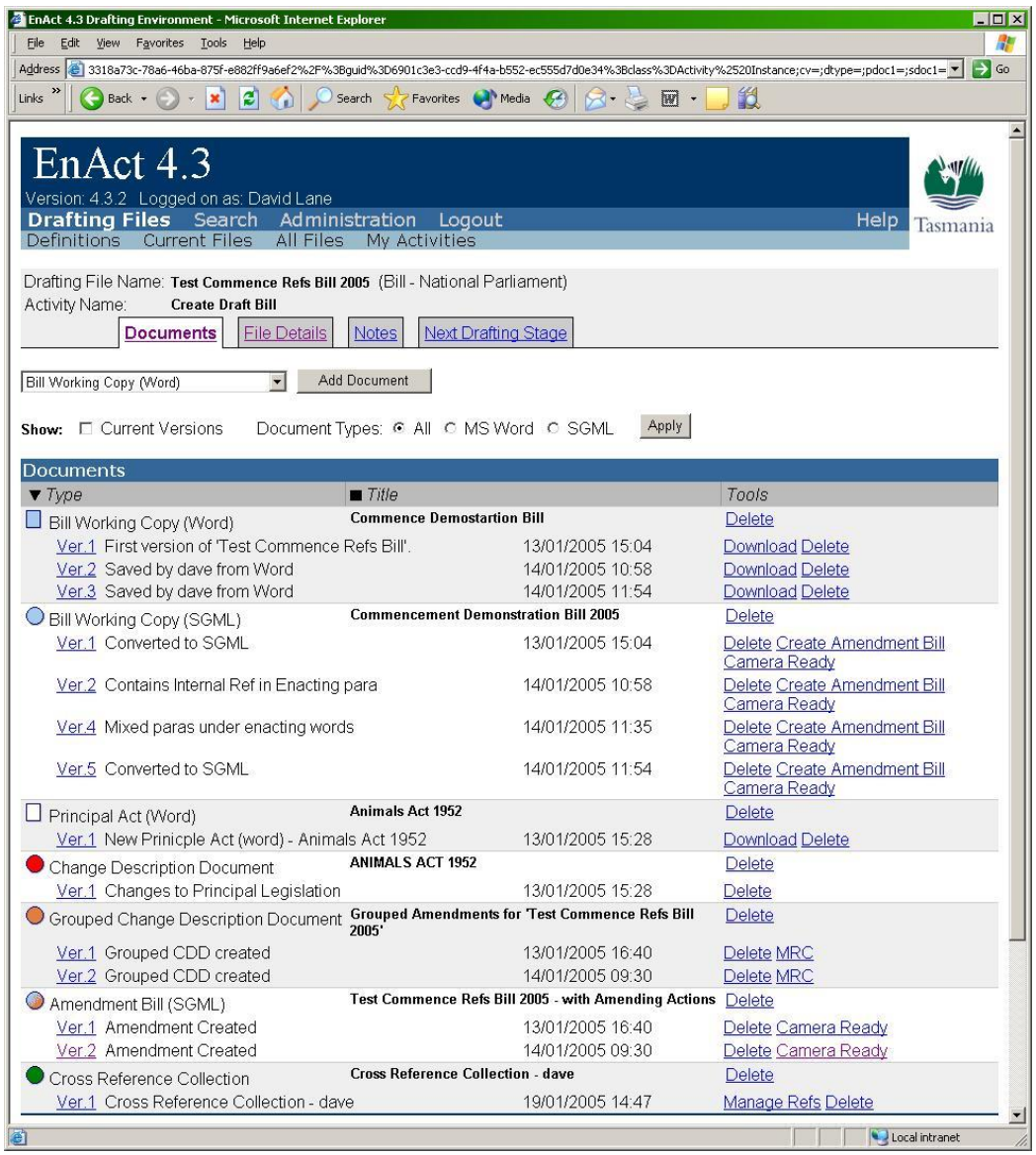
[Figure 2] User worklist in the EnAct system

Figure 3 shows a page for updating the metadata for a particular Bill project. What is displayed on this page depends on the stage of the project with different fields either locked or made available to complete as the project progresses. The project shown is a regulation at an advanced stage being prepared to send to the government printer (Printing Authority of Tasmania).



[Figure 3] Updating project metadata in the EnAct system

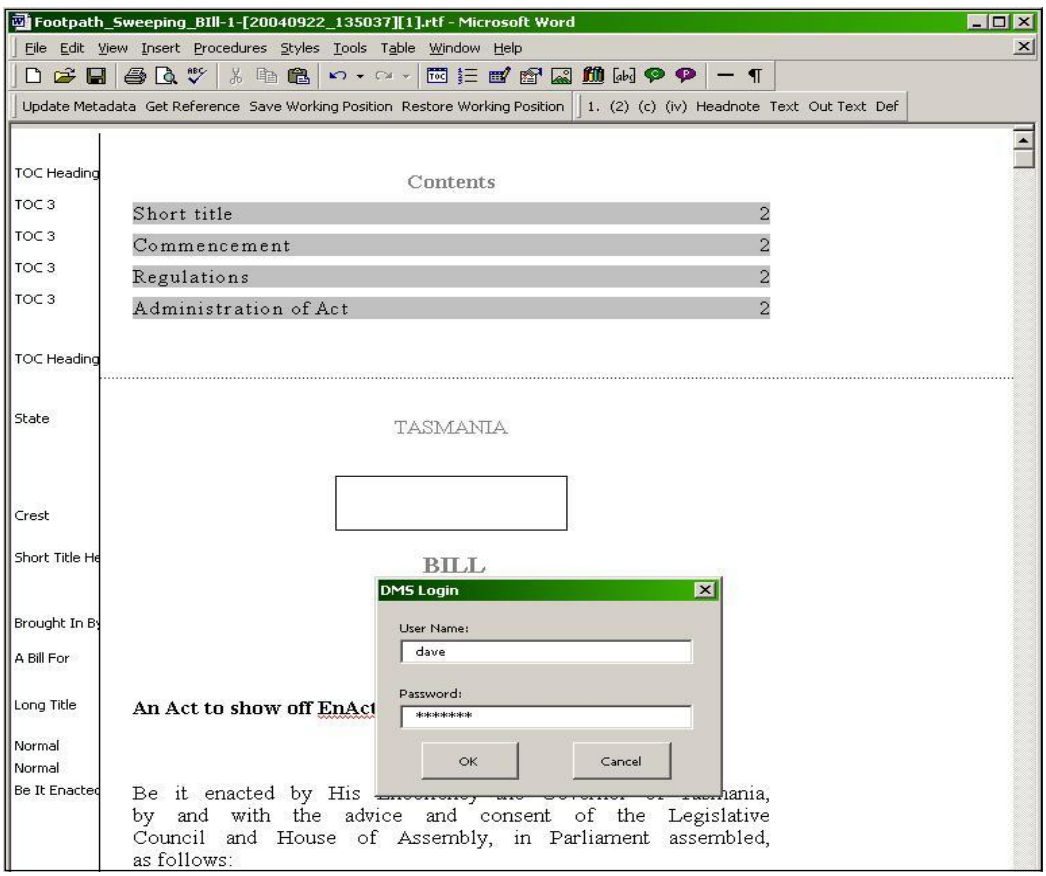
Figure 4 shows a page displaying the different documents and versions of documents associated with a particular drafting task. This page shows both Microsoft Word and SGML versions of the draft amending Bill together with the change-tracked Principal Act (the Act being amended). The document types peculiar to the EnAct system are the grouped and raw Change Description Documents (used to manage automatic generation of amendment wording) and the cross-reference collection (used to manage the creation and insertion of standardized cross-references into legislative documents).



[Figure 4] Display of document versions in a Bill project in the EnAct system

Ⅳ Example custom tools for authoring and editing

It is rare for a drafting office to take an Editing Tool without customizing it to adapt to the house style and processes of the drafting office. For XML structured authoring environments, this is necessary to provide a formatted view (WYSIWYG[8]) of the documents while drafting. For word processor environments, this may extend to defining a set of styles to use. For both, this may include a variety of custom macros and functions from globally replacing acronyms to popping up the name of the person allocated on today's tea roster. We focus now on some specific customizations to support the legislative drafting process.



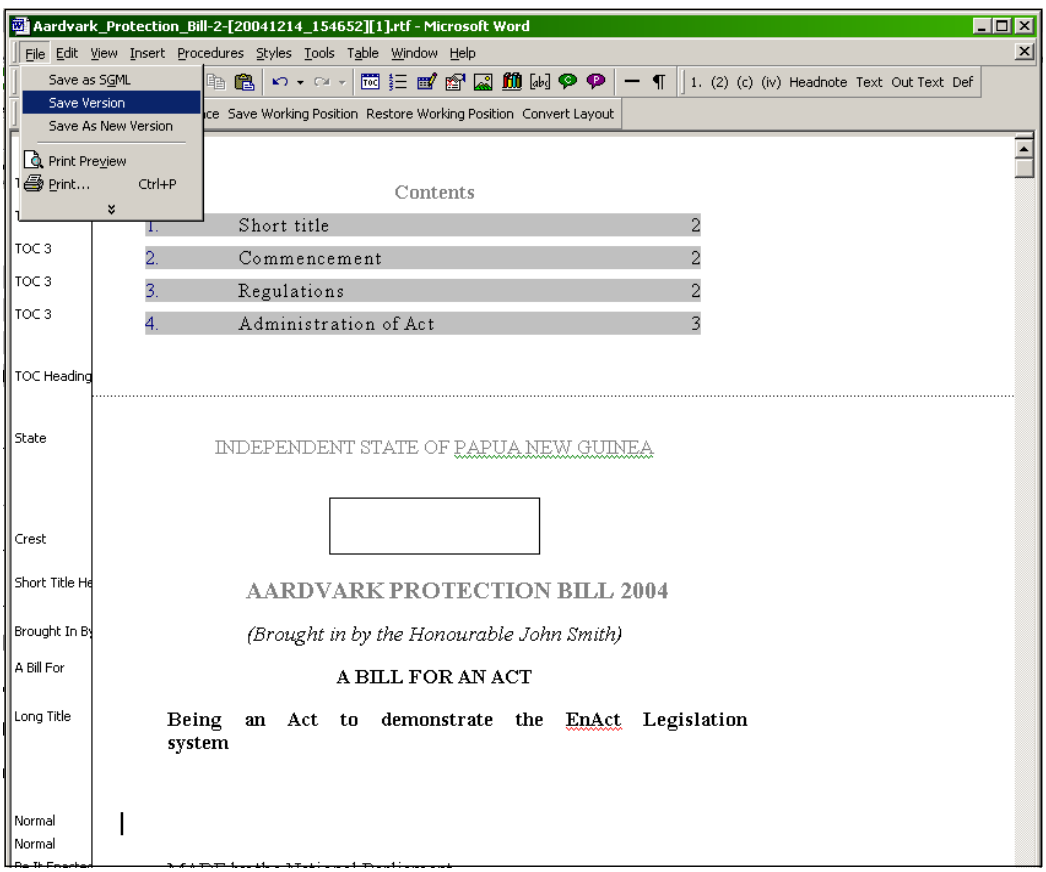
[Figure 5] Login from Microsoft Word in EnAct system

8) WYSIWYG: What You See Is What You Get

Editor customization for version management

Where a Document Management System is involved, it is desirable to access some or all of the DMS functionality directly from the Editing Tool. Figure 5 shows the login screen which is displayed the first time a DMS function is accessed from the EnAct Editing Tool (Microsoft Word).

Figure 6 shows some customized Save Options in the EnAct system. Note the non-standard "Save as SGML" which automatically converts the strongly styled Microsoft Word document into XML (SGML) conforming to a custom EnAct document type definition (DTD), and the more generic "Save Version" and "Save as New Version" functions which overwrite the latest version or create a new version respectively.



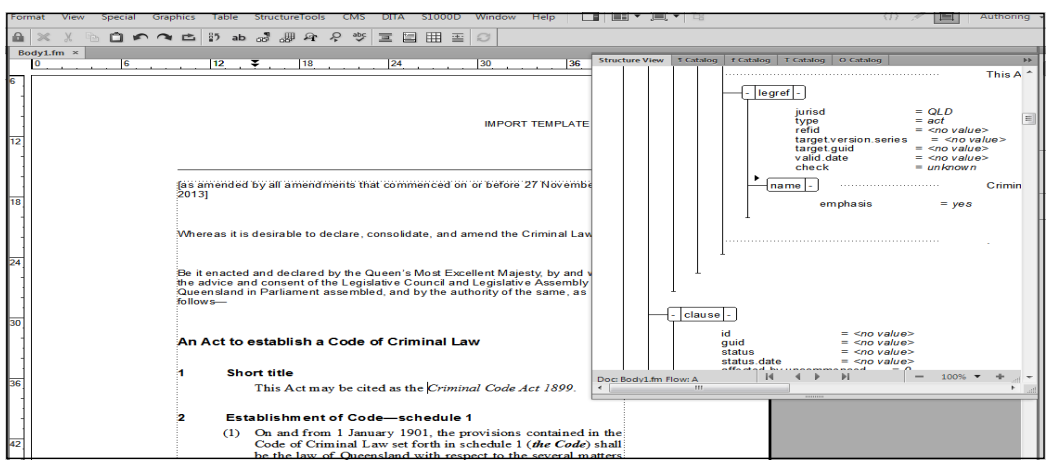
[Figure 6] Save options from Microsoft Word in EnAct system

Where a chosen Editing Tool also supports the generation of other formats (such as Adobe FrameMaker which produces PDF and XML versions in addition to the native FrameMaker format), additional functions can be added to generate special versions with particular templates for introduction, consolidation of amendments on the floor, and Assent and similar.

We have experienced significant success using the HTTP- and XML-based protocol, SOAP, as a generic integration framework as it is accessible from a wide variety of programming languages and has been used by us to integrate Adobe FrameMaker, Microsoft Word, XMetal, and WordPerfect®[xii] platforms with little or no change to the basic API[9] for the particular Editing Tool or supported formats.

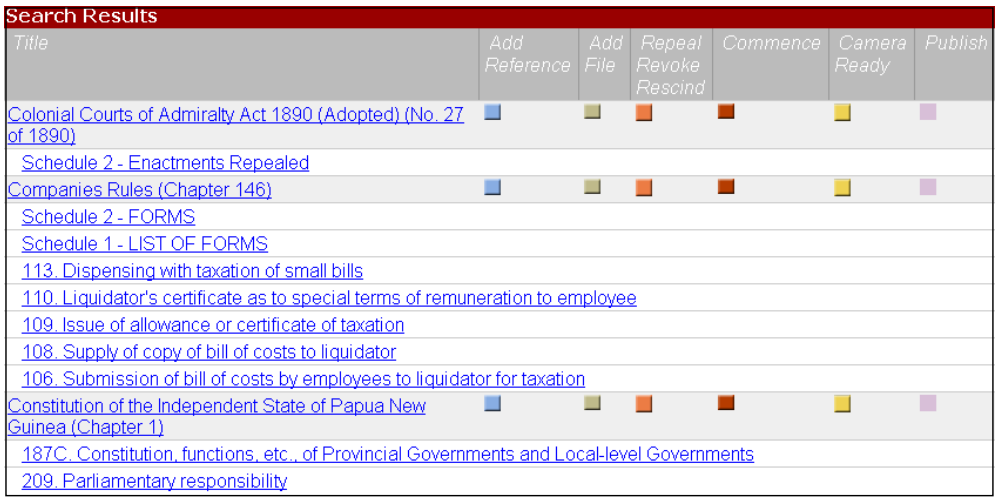
Editor customization for specialized link management

The normal method for managing cross-reference markup in an XML environment is simply to populate any identifier attributes manually so that hypertext links can connect the cross-reference wording to the appropriate target document or provision. Figure 7 shows manual insertion of cross-reference markup in Adobe FrameMaker which is the default way of managing cross-references in most structured authoring environments without any customization (particularly references to other documents or provisions within other documents).



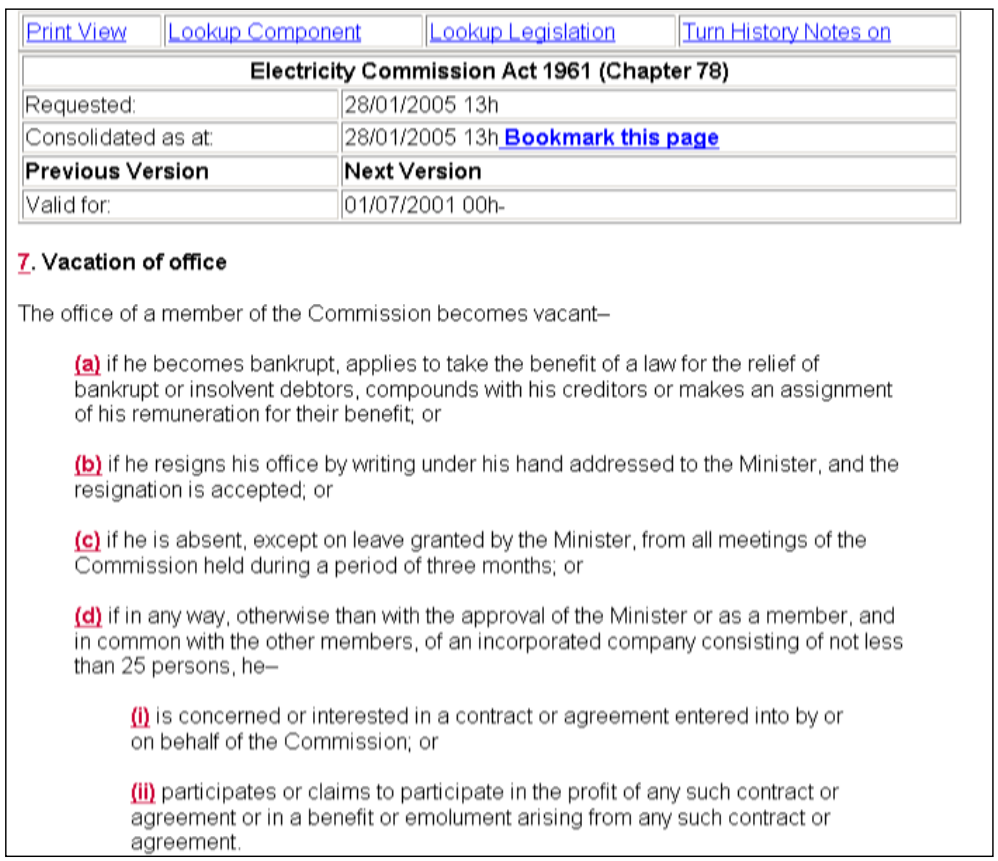
[Figure 7] Manual insertion of cross-reference markup

9) API: Application Program Interface


Search Results

Title	Add Reference	Add File	Repeal Revoke Rescind	Commence	Camera Ready	Publish
Colonial Courts of Admiralty Act 1890 (Adopted) (No. 27 of 1890)						
Schedule 2 - Enactments Repealed						
Companies Rules (Chapter 146)						
Schedule 2 - FORMS						
Schedule 1 - LIST OF FORMS						
113. Dispensing with taxation of small bills						
110. Liquidator's certificate as to special terms of remuneration to employee						
109. Issue of allowance or certificate of taxation						
108. Supply of copy of bill of costs to liquidator						
106. Submission of bill of costs by employees to liquidator for taxation						
Constitution of the Independent State of Papua New Guinea (Chapter 1)						
187C. Constitution, functions, etc., of Provincial Governments and Local-level Governments						
209. Parliamentary responsibility						

[Figure 8] Using the legislation repository for link target identification in EnAct


Print View | Lookup Component | Lookup Legislation | Turn History Notes on

Electricity Commission Act 1961 (Chapter 78)

Requested:	28/01/2005 13h
Consolidated as at:	28/01/2005 13h **Bookmark this page**
Previous Version	**Next Version**
Valid for:	01/07/2001 00h-

7. Vacation of office

The office of a member of the Commission becomes vacant–

(a) if he becomes bankrupt, applies to take the benefit of a law for the relief of bankrupt or insolvent debtors, compounds with his creditors or makes an assignment of his remuneration for their benefit; or

(b) if he resigns his office by writing under his hand addressed to the Minister, and the resignation is accepted; or

(c) if he is absent, except on leave granted by the Minister, from all meetings of the Commission held during a period of three months; or

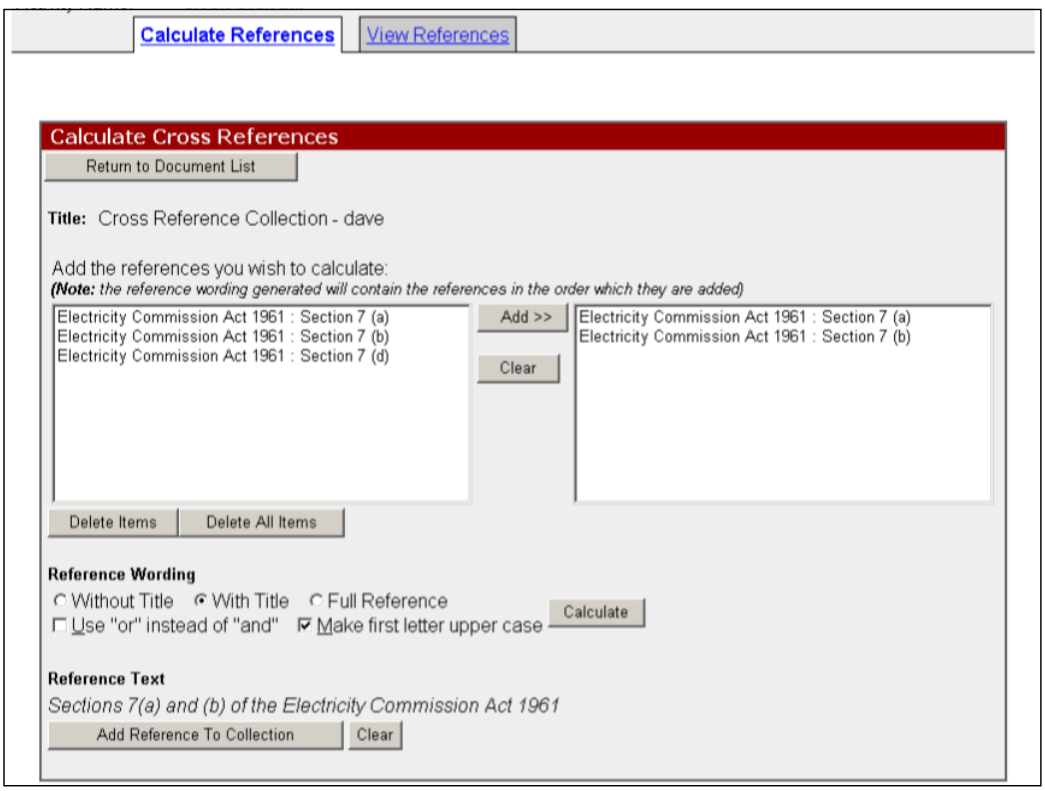
(d) if in any way, otherwise than with the approval of the Minister or as a member, and in common with the other members, of an incorporated company consisting of not less than 25 persons, he–

(i) is concerned or interested in a contract or agreement entered into by or on behalf of the Commission; or

(ii) participates or claims to participate in the profit of any such contract or agreement or in a benefit or emolument arising from any such contract or agreement.

[Figure 9] Identification of individual provisions for link generation

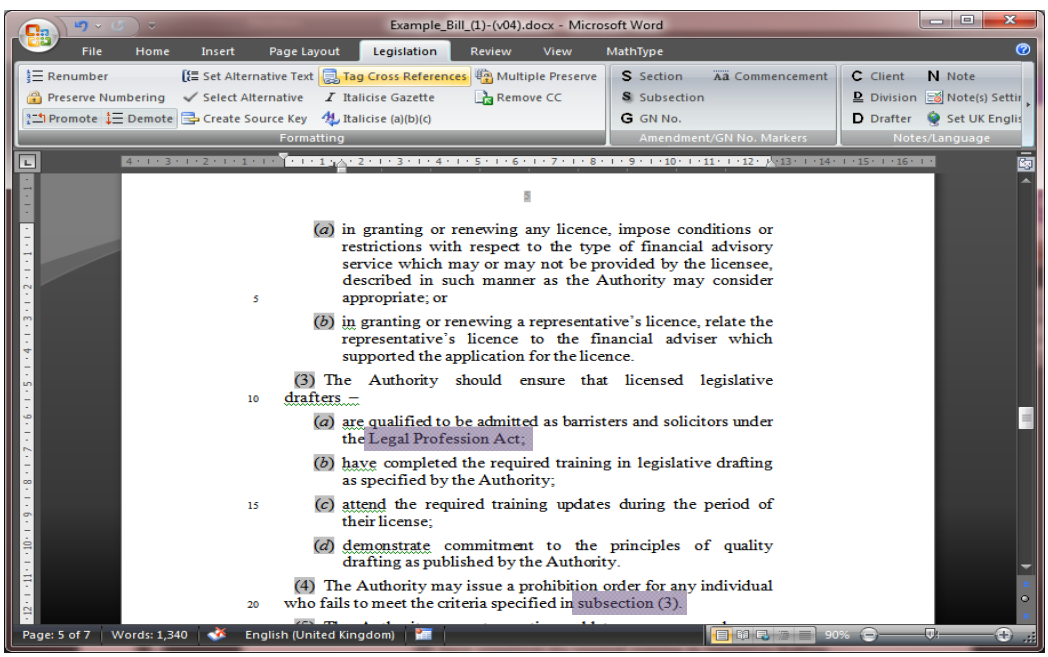
By contrast, the EnAct system makes use of the legislation repository with a custom web interface to identify link targets. Figure 8 shows a search result in the EnAct legislation collection. When a title or provision is selected, Figure 9 allows the user to display that provision. Clicking on the red number adds that particular target to a list of target provisions maintained for each project. Figure 10 shows the result of clicking on (a), (b) and (d) in Figure 9 to select the particular targets below the section level. The "Add >>" button allows a user to select candidates for inclusion in a particular set of reference wording. Items under "Reference Wording" allow the user to customize the wording generated including or excluding the title of the Act, choosing the long or short form of reference wording, and selecting the connecting conjunction. The "Add Reference To Collection" button adds that wording to the collection so it can be inserted complete with all identifiers into a draft document.



[Figure 10] Turning identified provisions into standardized reference wording with markup

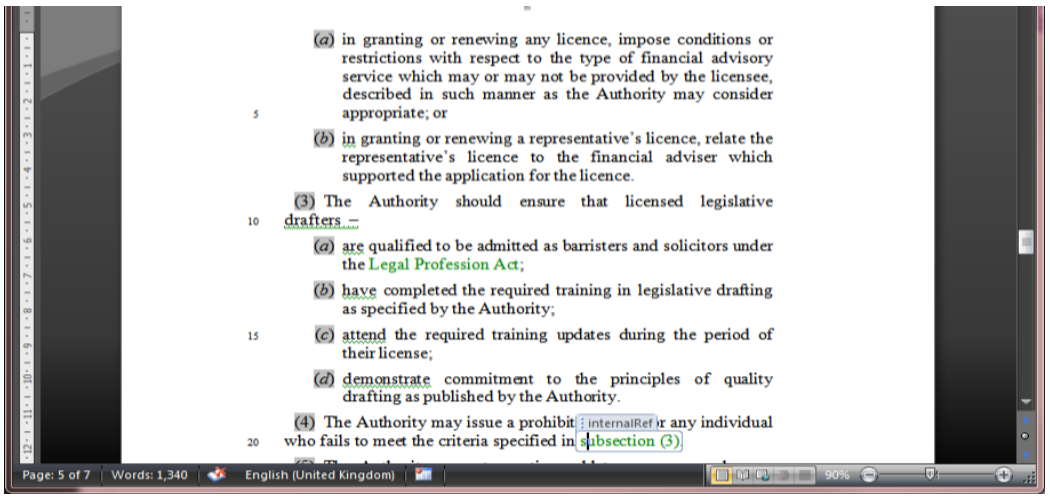
While this approach allows for standardized wording and insertion of required identifiers without exposing the user to the detail of the underlying markup, there is still some manual effort required.

An alternative approach has been used in Singapore and Queensland,[10] where the users simply type the wording for the cross-reference without disrupting the flow of drafting – see Figure 11. I have highlighted the text in this fragment that corresponds to cross-references. A user then clicks the "Tag Cross References" button which sends a copy of the document to the server which searches for known document titles and patterns of wording that reflect reference wording and attempts to resolve them against the collection, populating all the necessary identifier attributes and inserting them into the draft document (in this case a Microsoft Word document using content controls) and styling the references green (for the drafting only) to show that the reference has been successfully resolved – see Figure 12.



[Figure 11] Automatically recognizing reference text – part 1

10) Similar approaches are outlined in the work of AustLII (e.g., Greenleaf, G *et al.* "Public access to law via the internet: the Australasian Legal Information Institute" (1995) 6(1) *Journal of Law, Information and* Science 49) and NIR (e.g., Biagioli, C *et al.* "Automatic semantics extraction in law documents" in *Proceedings of the 10th International Conference on Artificial Intelligence and Law*. ACM, 2005).



[Figure 12] Automatically recognizing reference text – part 2

While this approach cannot be perfectly accurate (the most common error is marking references to provisions in other documents as internal references), a success rate above 99% has been experienced in practice and either of the approaches described above can be used to correct the small percentage of incorrectly identified references manually. The success rate experienced in a particular jurisdiction will depend on the regularity of the cross-reference wording used; however, accuracy below 90% would be unlikely.

Because it is relatively easy to markup cross-references in this way, including internal cross-references (references to other provisions within the current document), this approach can be applied repeatedly throughout the drafting process with little imposition on the system users. It also delivers a substantial benefit to the drafting office because draft Bills (and sometimes draft Regulations) typically don't get renumbered during their passage through the legislature. New provisions inserted through committee amendments and on the floor are numbered like inserted provisions in enacted legislation. It is only for the third reading, or assent copies, that the provisions are renumbered. If the references are correctly tagged in this way, automatic renumbering can also resolve the reference wording correctly to ensure that the appropriate target provision is represented in the reference wording.

This is an example of a "win-win" automation tool for the drafting office and external

consumers of legislation as it provides added utility to both external and internal users, and imposes little, if any, additional burden on the internal users.

Editor customization for automatic amendment wording

The ability to automatically generate reference wording is a necessary prerequisite for the more difficult task of automating amendment wording generation. Since at least 80% of legislation is amending legislation (legislation that changes existing law rather than creates new substantive law), improved tools for drafting amendments have larger potential for productivity gains than tools to support substantive law drafting.

Since the goal of amendment drafting is to produce a cohesive and correct law as amended (the consolidation) and the actual wording describing the amendment is a fleeting artifact, it makes sense to start the amendment drafting process by marking the proposed changes on the appropriate version of the Principal Act or regulation. Figure 13 shows a provision of a Principal Act ready for amendment.

Obligations of licensees

74F. A licensee -

(a) must not make bets except in accordance with the licence; and

(b) must not, as a licensee, make bets on horse races or greyhound races; and

(c) must not conduct betting or business under the licence on Good Friday or Christmas day or at other times notified to the licensee in writing by the Supervising Agency; and

(d) must, for verification purposes, make a recording of all bets made by telephone under the licence in such manner as the Supervising Agency from time to time directs, the costs of such verification to be at the expense of the licensee; and

(e) must keep such betting, accounting and other records in respect of the licensee's betting and business under the licence as the Supervising Agency from time to time directs; and

(f) must furnish to the Supervising Agency such returns in respect of the licensee's betting and business under the licence as the Supervising Agency from time to time directs.

[Figure 13] Principal Act before amendment

Figure 14 shows two amendments marked on this Principal Act, the first a textual amendment substituting old wording for new wording, the second is a substitution of a whole

XML element–a paragraph.

Figure 15 shows a screen shot of the EnAct system using custom buttons and styles to mark up changes in Microsoft Word. The current version deployed in Singapore uses standard Microsoft Word change tracking to identify the inserted and deleted text and provisions.

Figure 16 shows the process by which this markup is converted into amendment wording. First the XML version of the Act to be amended is checked out of the repository. The system turns this into a Microsoft Word document for editing (although this could be done using a structured XML editor and custom or tool-based change markup). The user then marks up the proposed changes to the Principal Act either using custom buttons and styles or using the native change tracking. The system then generates a Change Description Document (CDD) from this change tracking. The CDD is an XML encapsulation of all the information needed to reproduce that amendment when applied to that XML version of the Principal Act including all of the necessary context required to generate accurate reference wording. The CDD is then used to generate an XML amendment document (which can be exported as a Microsoft Word document for further editing if required).

Obligations of licensees

74F. A licensee -

(a) must not make bets except in accordance with the licence; and

(b) must not, as a licensee, make bets on horse races or greyhound races; and

(c) must not conduct betting or business under the licence ~~on Good Friday or Christmas day or at other times~~ at any time notified to the licensee in writing by the Supervising Agency; and

~~(d) must, for verification purposes, make a recording of all bets made by telephone under the licence in such manner as the Supervising Agency from time to time directs, the costs of such verification to be at the expense of the licensee; and~~

(d) in respect of any bet made other than in person, must electronically record the bet for verification purposes -

(i) by a device approved by the Supervising Agency; and

(ii) in a manner directed by the Supervising Agency; and

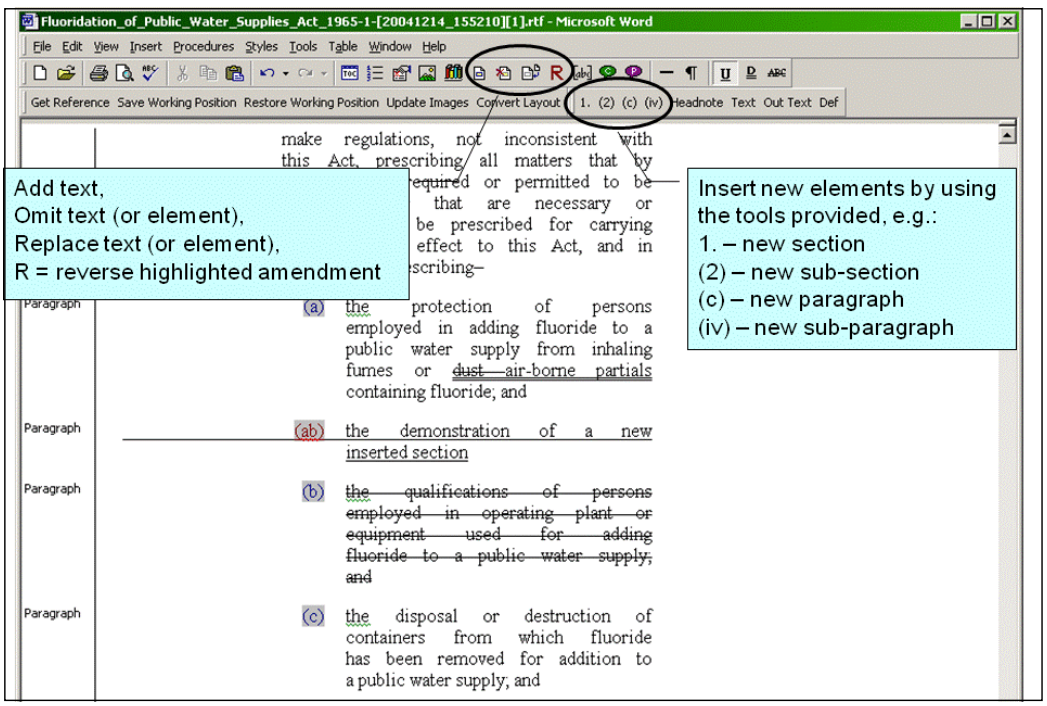
(e) must keep such betting, accounting and other records in respect of the licensee's betting and business under the licence as the Supervising Agency from time to time directs; and

(f) must furnish to the Supervising Agency such returns in respect of the licensee's betting and business under the licence as the Supervising Agency from time to time directs.

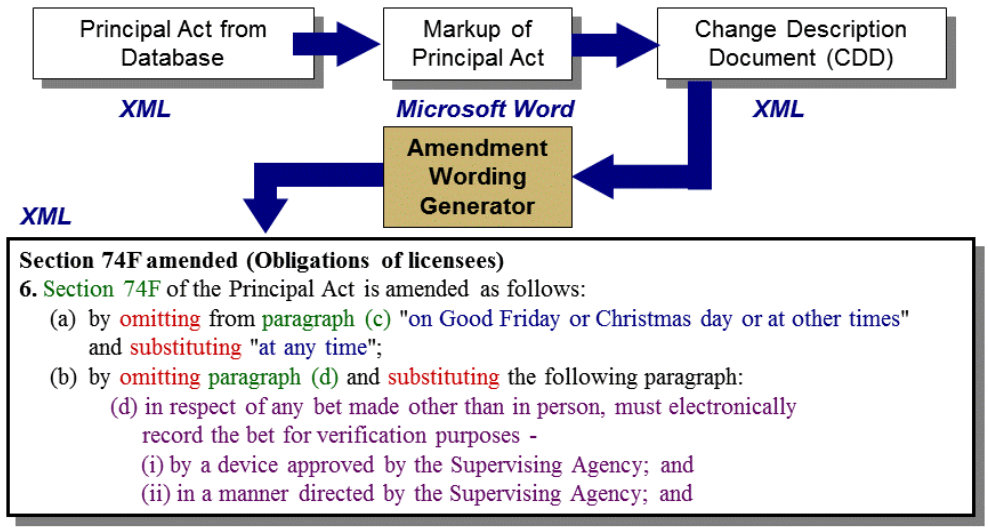
[Figure 14] Principal Act with change-tracked amendments

This process has the advantage that the user focuses on the wording of the resulting consolidation rather than the fleeting amendment wording. It also generates very standardized amendment wording and reference markup. In addition, this process aids external users in interpreting the amendments consistently (although, as will be seen in the next section, that becomes of less importance).

In Tasmania, users mark the changes using direct styling of strike-through and underline but in Singapore users make use of the change-tracking capabilities of Microsoft Word (these weren't available with the required capability in the versions of Microsoft Word that were used for the first version of the Tasmanian EnAct system).



[Figure 15] Marking changes using custom buttons in EnAct



[Figure 16] Generating the amendment wording

V Example custom tools for publishing

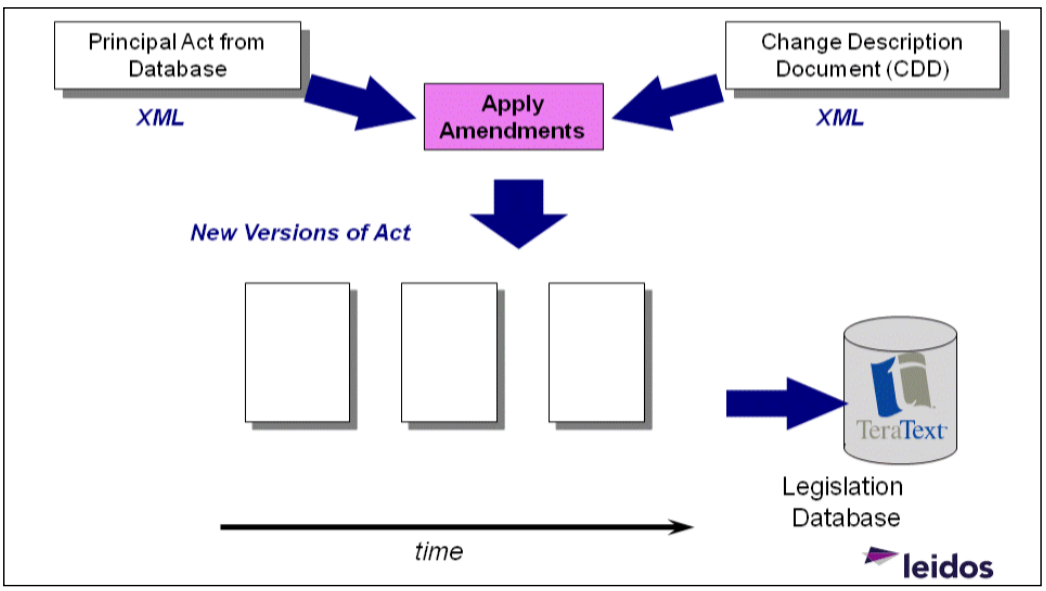
Automatically applying amendments

If the approach above is applied to generate amendment wording, the Change Description Document, or CDD, can be associated with the amending Bill as it progresses through the enactment stages and, once the Bill becomes an Act and the commencement details are known, the CDD can be used to automate the application of the amendment against the previous version of the Principal Act to create one or more new versions of the Act.

This new version of the Act can then be saved in the document repository and the appropriate web objects created to support point-in-time access to the different versions of the legislatio n.[11] Although not applied commercially, it is also possible to extract these CDDs from an amending Act created without using the automatic amendment generation.[12]

11) Arnold-Moore, T "Automatic generation of amendment legislation" in *Proceedings of the 6th International Conference on Artificial Intelligence and Law*. ACM, 1997.

12) Arnold-Moore, T "Automatically processing amendments to legislation" in *Proceedings of the 5th*



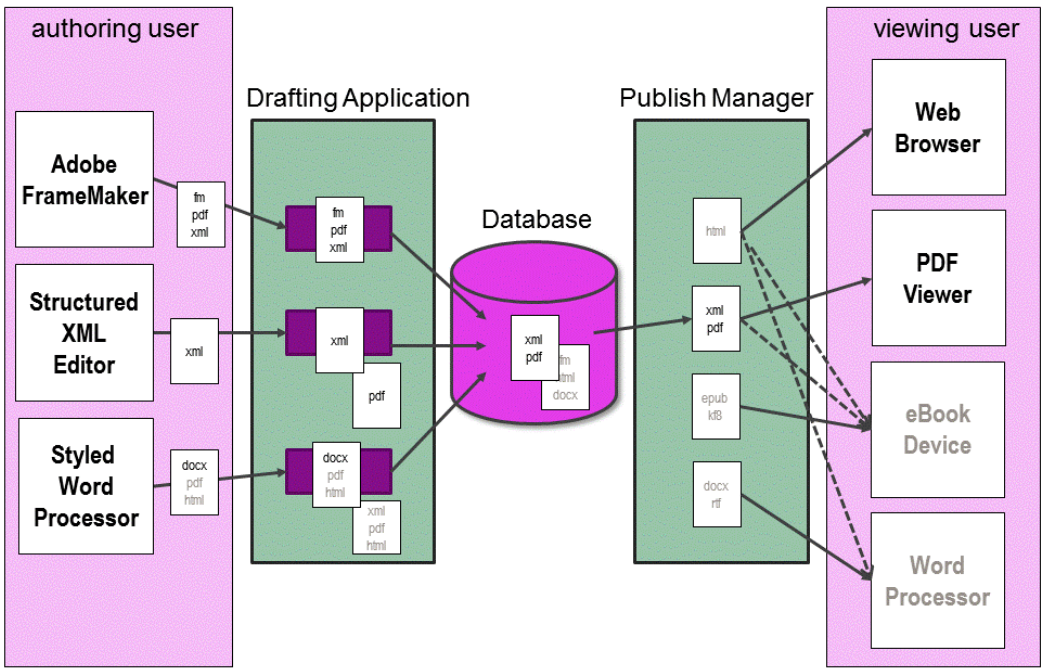
[Figure 17] Applying marked changes to produce new versions of the Principal Act

Managing and creating multiple versions

Figure 18 shows a diagram of the various formats that might be useful and how they might be generated. Working back from the desired public website, users want to be able to access HTML (and possibly XML) using standard web browsers (e.g. Safari®[xiii], Chrome™[xiv] and Internet Explorer®). They might also wish to download a PDF version of the relevant legislation to faithfully reproduce the printed page. For reuse of the content, a Microsoft Word version (whether RTF or OOXML/docx) is also desirable and users of eBook readers have better functionality from ePub℠[xv] or similar eBook formats.

Where the Editing Tool is Adobe FrameMaker, the Editing Tool can generate FrameMaker, PDF, and XML formats directly which can be saved and managed by the drafting Application. If necessary, the related Adobe FrameMaker Publishing Server can generate Microsoft Word (RTF) and ePub versions from the FrameMaker version without user intervention or stylesheets can be created to generate these from the XML automatically.

International Conference on Artificial Intelligence and Law. ACM, 1995.



[Figure 18] Managing and creating various formats for publishing

Where the Editing Tool is a standard structured editor (such as <oXygen/>, ArborText Editor, or XMetaL), the format saved from the Editing Tool is custom XML. Generating PDF for proof-reading and enactment processes requires a separate Render Service. In this circumstance, Adobe FrameMaker document types would be unlikely to be provided but the other formats could be generated automatically by applying stylesheets to the XML.

Where the Editing Tool is Microsoft Word, the format saved from the Editing Tool is Microsoft OOXML format (or RTF). This format can be converted to custom XML (or SGML) providing the styles are tightly controlled and managed within the Microsoft Word environment. In the absence of custom XML, Microsoft Word can generate PDF and HTML directly. If custom XML is used, ePub, and possibly PDF, can be generated using stylesheets and a separate Render Service respectively.

All of these processes require considerable customization – Adobe FrameMaker requires read-write rules to be defined specific to your custom XML DTD, the other structured editors

require stylesheets to specify the formatting within the Editing Tool as well as for formatting the custom XML to the output formats. The Microsoft Word approach requires considerable styling and macros to make it easier for users to generate content that can be easily converted to a custom XML DTD as well as configuration of the conversion utility. Stylesheets are required for rendering that XML to the other formats. If native Microsoft Word is the source format, there is still processing needed of the HTML that is produced by Microsoft Word to make it usable in a generic public website.

Except for the native Microsoft Word source model, each of these models assumes a considerable data migration phase to get the legislation corpus into a custom XML DTD. The native Microsoft Word source model may require some minimal document cleanup but there is likely to be a more comprehensive on-going data migration cost every time a new version of Microsoft Word is deployed in the drafting office.

However, each of these different models has been successfully deployed in a drafting office to manage the publication of legislation and related materials with sufficient automation to ensure that the effects can be achieved with minimal additional effort by the drafting office.

Ⅵ Conclusion

In this paper, we have demonstrated that the unique nature of legislative documents and the processes by which they are created establishes an environment in which special-purpose tools deliver real benefit to both the drafting office that prepares them and the external consumers of legislative documents and related material.

In 1799, Justice Lawrence argued for publishing the results of legislative deliberations more broadly:

> “The proceedings of Courts of Justice are daily published (...). The general advantage to the country in having these proceedings made public, more than counterbalances the inconveniences to the private persons whose conduct may be the subject of such proceedings. The same reasons

> also apply to the proceedings in Parliament: it is of advantage to the public, and even to the legislative bodies, that true accounts of their proceedings should be generally circulated." per Lawrence J, *R v Wright* (1799) 8 TR 293, 298: 101 ER 1396, 1399.

In 2000, New Zealand Parliamentary Counsel Office extended this argument suggesting a broad government obligation to make up-to-date legislation available on the web without cost:

> "Legislation confers rights, benefits, and privileges, and imposes obligations. It is a fundamental principle that ignorance of the law is no excuse for a failure to comply with the law. People cannot be expected to know their legal rights or comply with the law if the law is not easily accessible. It is for the State to ensure that it is. Technology has now advanced to a point where access to up-to-date legislation can be provided free via the Internet." New Zealand Parliamentary Counsel Office (2000) *Public Access to Legislation Project: Summary of business case*

The tools and techniques described in this paper demonstrate that governments can go well beyond this modest objective to deliver better and more effective tools to support the legal profession and improve the cost and efficiency of legislative and broader legal services to the community.

---

[i] Arbortext, and Arbortext Adapter are registered trademarks of PTC, Inc. in the United States and/or other countries.
[ii] Xmetal is a registered trademark of JustSystems Canada, Inc. in the United States and/or other countries.
[iii] Adobe, FrameMaker, the FrameMaker icon, and the PDF file icon are registered trademarks of Adobe Systems Incorporated in the United States and/or other countries.
[iv] 〈oXygen/〉 is a registered trademark of SC SyncRO Soft SRL in the United States and/or other countries.
[v] OpenOffice and SOLR are registered trademarks of The Apache Software Foundation in the United States and/or other countries.
[vi] Documentum is a registered trademark of EMC Corporation in the United States and/or other countries.
[vii] TeraText is a registered trademark of Leidos, Inc. in the United States and/or other countries.
[viii] Word launch icon, Internet Explorer, Internet Explorer icon, SharePoint, and SQL Server are trademarks or registered trademarks of the Microsoft Corporation in the United States and/or other countries.
[ix] JBOSS is a registered trademark of Red Hat, Inc. in the United States and/or other countries.
[x] Oracle, JAVA, and WebLogic are registered trademarks of Oracle International Corporation in the United States and/or other countries.
[xi] WebSphere and DB2 are registered trademarks of the International Business Machines Corporation in the United States and/or other countries.
[xii] WordPerfect is a registered trademark of the Corel Corporation in the United States and/or other countries.
[xiii] Safari is a registered trademark of Apple, Inc. in the United States and/or other countries.
[xiv] © 2016 Google Inc. All rights reserved. Chrome is a trademark of Google, Inc. in the United States and/or other countries.
[xv] ePUB and the ePUB logo are registered Service Marks of International Digital Publishing Forum in the United States and/or other countries.