

# RE-ENGINEERING ONLINE CASE LAW: LEGAL RESEARCH IN SINGAPORE

Ken Hwee Tan\*

## Abstract

This brief paper recounts the development of online legal research tools in Singapore, and describes the various entities involved. It then proceeds to focus on the case law databases that are available, and explains the plurality of backend systems that have proliferated over the years, with significant implications for the further development of these tools, and the quality of service that can be expected by users of the system.

It discusses the challenges that have been identified, and describes a possible scenario where, inter alia, the various databases are brought within a single overarching content management structure, and the promises and advantages of such an approach. Finally, it concludes by explaining the current state of play in this search for content nirvana.

The main purpose of this paper is to document our experiences and the challenges that we have identified for the near future, with a view towards sharing our experiences, and also benefiting from the experience of others who may have faced the same challenges and conquered them already.

## Background

### WHAT IS THE SINGAPORE ACADEMY OF LAW?

The Singapore Academy of Law was created in 1988 after the Singapore Academy of Law Act was passed by Parliament. The Academy was created to foster scholarship and learning of the law, and also to take charge of the dissemination of the law. To fulfil this role, the Academy is involved in continuing legal education, publishes the Singapore

---

\* Assistant Director, Singapore Academy of Law, and State Counsel, Attorney-General's Chambers, advising on public international law and technology law.

Academy of Law Journal, and manages Singapore's online legal research infrastructure.

## WHAT IS LAWNET?

More than ten years ago, some officers in the Attorney-General's Chambers, together with staff from the National Computer Board (as it was then called) came up with the idea of sharing some public sector databases with law firms. However, in order for various Government departments to open up their databases, they would each have to arrange for connectivity, billing, helpdesk and authentication mechanisms. This was a major stumbling block. The solution was to utilise a "value-added network vendor" to act as a one stop clearinghouse for these services. Singapore Network Services Pty Ltd was asked to provide these services. The "umbrella" or "glue" holding the various services together would be the authentication, billing, and helpdesk services provided by SNS Pty Ltd. The initial services made available to members of the public included:

- company registration information from the Registry of Companies and Businesses;
- full text of statutes, and then subsidiary legislation, from the Attorney-General's Chambers STATUS database; and
- full text of cases, in collaboration with Butterworths Asia, from the Attorney-General's Chambers.

LawNet was therefore created, with government support and funding for its initial five year period, as a project managed out of the Attorney-General's Chambers. It was however expected to be self financing after five years. This, together with the need for LawNet to grow and adapt to new technologies with as much agility as possible, led to the transfer of the management of LawNet from the Attorney-General's Chambers to the Academy of Law.

LawNet, as conceptualised by its founders, covered various modules: legal research, litigation, corporate law, intellectual property, and conveyancing. The rest of this paper focuses on the legal research module, and specifically, the case law components in this module.

## WHAT IS THE LEGAL WORKBENCH?

LawNet was launched at a time when the internet was still the playground of academics and research institutes. It initially comprised various databases and host systems, all tied together from an authenti-

cation, billing and support point of view, but each disparate and requiring a different user interface insofar as function keys and commands were concerned. With the advent of the internet, it was only a question of time before LawNet took advantage of this new technology. In 1998 LawNet legal research fully embraced the internet platform through the launch of Singapore's second generation online legal research tool, the Legal Workbench ("LWB").

Major case law and legislation databases were moved to an internet platform, with legislation seeing its third incarnation (the first version was a STATUS database, the second a simple HTML collection called Legis.Online, and the current system, the Versioned Legislation Database system, is based on SGML). With this move came a decision to radically change the charging paradigm for the legal research module. Instead of charging on a time-based and a per search or per record basis for the terminal-based access, users were asked to pay on a flat fee unlimited usage basis. In addition, to ensure that large firms did not yield to the temptation of paying for a few subscriptions and sharing user IDs amongst their lawyers, charges were levied on the basis of the total number of lawyers in each firm. Law firms are not able to specify a lower number of users (eg for those in the litigation department only) but the quid pro quo is that the per user charge is significantly discounted for each additional user.

This new charging paradigm was predicated by the notions that time-based charging does not work well for internet services, and that charging for CPU time or for each search also seemed anachronistic. These notions are not sacrosanct and in fact, since the LWB was launched the *technical* possibility of time-based charging has reappeared, and has been implemented for ad hoc users who pay for access by credit card or stored value cashcards because they do not wish to maintain a fixed fee monthly subscription commitment.

What has not changed is perhaps the most important reason why there was a major effort to convince subscribers to take on fixed fee subscriptions. Legal research, especially online legal research through the use of search tools, is most effective when it is an iterative process. The trial and error involved in constructing a search query, executing it, and refining it does not augur well for charging on a per search basis.

At this time, more than 70 per cent of the legal profession in Singapore (defined to be persons holding practising certificates), maintain, through their law firms, flat fee monthly subscriptions to the LWB.

## WHAT DATABASES ARE IN THE LWB?

Since its launch in 1998, the LWB set of databases has grown. Additional services have been added, and in each case, although there was an option to designate new services as “premium” services attracting additional fees, the additional services were added onto the “basic” set of services included within the basket of services available “by default” in the LWB.

Focusing, for the purposes of this paper, on case law and case law-related databases, what follows is a description of the various services within LWB, with a brief description of the technical underpinnings and the content creation chain pertaining to each service.

### The Legal Prospector

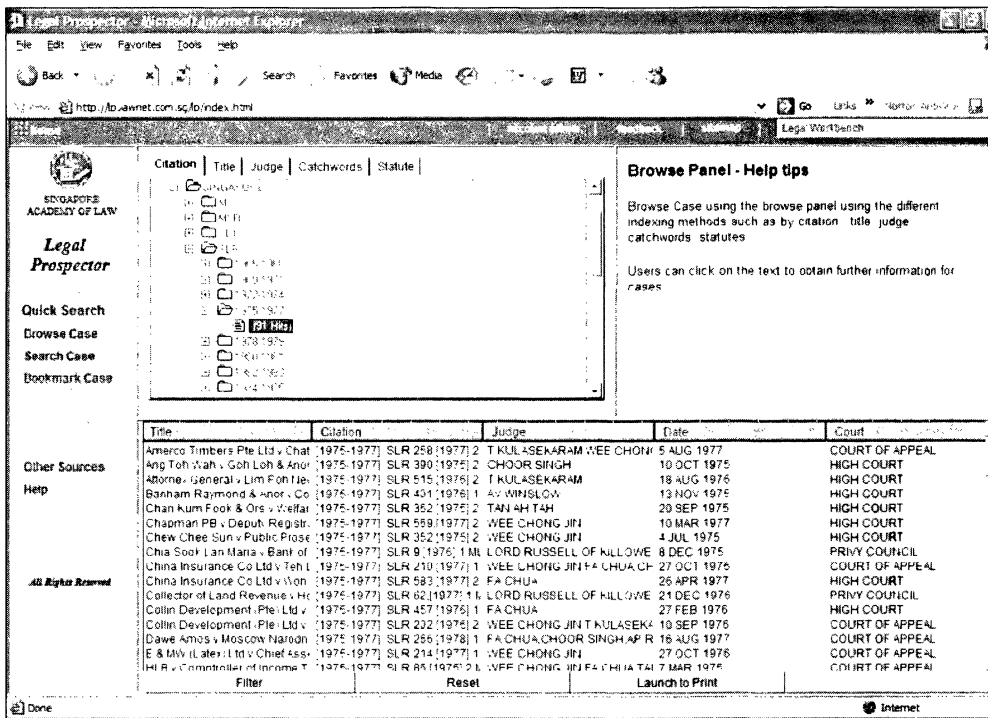


Figure 1: Legal Prospector (advanced interface)

Legal Prospector is the main case law database. It contains the full text of reported decisions, in collaboration with Butterworths Asia. Butterworths Asia, pursuant to a memorandum of understanding signed with the Academy, is authorised to produce the Singapore Law Reports. Butterworths Asia obtains the soft copy of raw judgments from the Academy, and when selected for reporting, prepares the reported version. It uses SGML in its backend processes, and uploads the SGML files to the LawNet computers. A rich set of meta-data components is required in order to provide for the flexible search and navigation features in Legal Prospector.

In 2001, as part of regular product development, Butterworths Asia was asked to provide the actual PDF of reported decisions, which are now available for download, for those users who wish to have a soft copy which, for all intents and purposes, is identical to the printed hard copy of the law reports.

On a technical level, Legal Prospector is relatively complex. The case meta-data is stored in SQL database tables, using Microsoft SQL Server 6.0. The full text and the PDF files are stored in designated file directories of a Microsoft Internet Information Server. Full text searches are carried out against the HTML directories and Microsoft Index Server. A special tool handles the upload of the SGML files from Butterworths Asia and parses it for the relevant meta-data to populate the SQL database tables. This tool is highly dependant on the precise Document Type Definition (DTD) for the SGML files, and any change in DTD can result in a need to rework the upload tool, or alternatively to down convert the document to match the earlier DTD.

## The Academy Digest

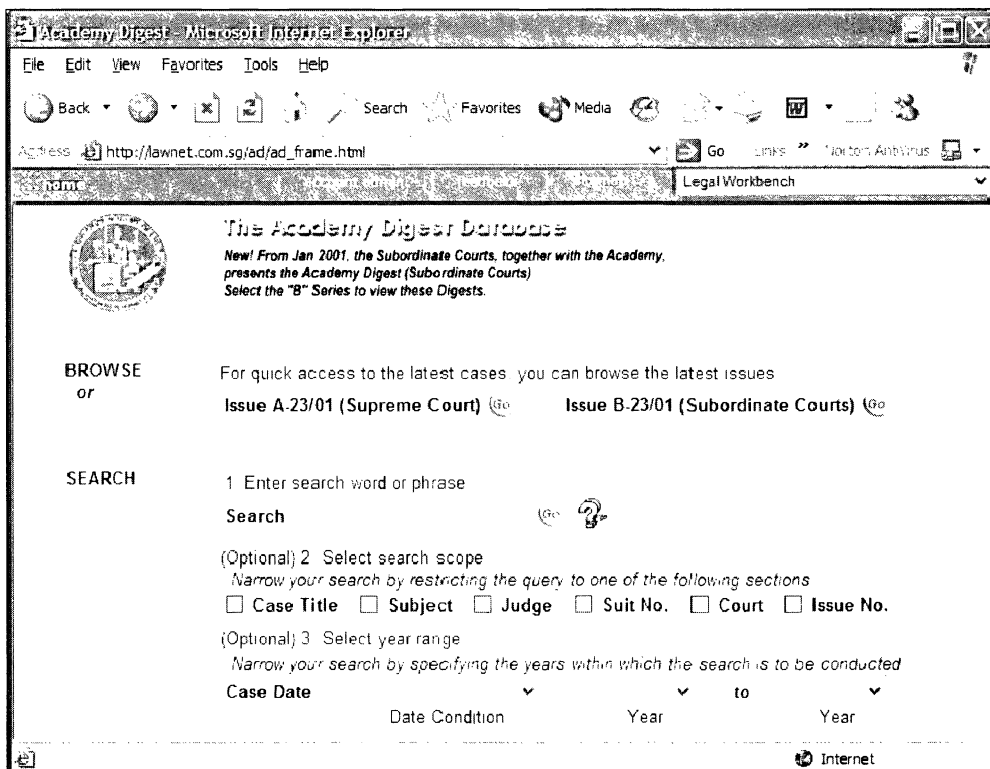


Figure 2: The Academy Digest

The Academy Digest is a series of case digests intended to provide brief summary information on all High Court and Court of Appeal cases, as quickly as possible after judgments are handed down, and in advance of the formal law reports. The digests started as a fax service for members of the Academy and other interested persons. In 1998, with the launch of the LWB, the Academy Digest subsystem was brought online within LWB so that users can search for summary information online, based on key fields such as the judge, case name or issue of the digest issue number. It is still available as a fax service.

The Academy produces the fax version of the digest, and staff in the LawNet Secretariat ensure that it is uploaded in properly formatted HTML files to the LawNet servers. On a technical level, the digest operates as carefully structured HTML files on a Microsoft Internet Information Server, indexed with Microsoft Index Server and

searched through search forms which make use of the meta-tags stored within the HTML files.

In 2001, the Academy Digest was enhanced to include digests of all judgments rendered by the Subordinate Courts, in addition to those rendered by the High Court and the Court of Appeal.

## The Unreported Judgments Database

**Unreported Judgments Database**  
New! From Jan 2001. Judgments from the Subordinate Courts are included  
LawNet thanks the Subordinate Courts for this valuable contribution

- Please specify whether you wish to see unreported judgments from
  - All Available Reports
  - Subordinate Courts Only
  - High Court Only
  - Court of Appeal Only
  - C A + H C
- Enter search word or phrase
- Select search scope (optional)
  - Anywhere in the judgment
  - Only search within --
    - Case Title
    - Subject
    - Sur: No
- Select year range (optional)
 

Date Condition    Year    to    Year

**Search**    **Help**

Figure 3: Unreported Judgments

With the launch of LWB, the Courts also made available unreported judgments. Before this service was available, access to unreported judgments was difficult and troublesome. Lawyers would have to track down hard copies at the Supreme Court or Subordinate Courts library, and they would have to know the actual judgment date or case number. With the online unreported judgments service (“URJ”), users could carry out full text searches against the judgments, without leaving their office.

Technically, URJ is relatively simple. Like the Academy Digest, it

relies on directories of properly meta-tagged HTML files. Where workflow is concerned, decisions are received from the Courts by the LawNet Secretariat. A secretary converts the files into HTML, making sure that any inconsistencies in formatting and paragraph numbering are standardised, and uploading the files to the LawNet servers.

URJ was enhanced in 2001 by including unreported judgments handed down in the Subordinate Courts, and by improving and streamlining the workflow processes that go into production.

### Cases Index

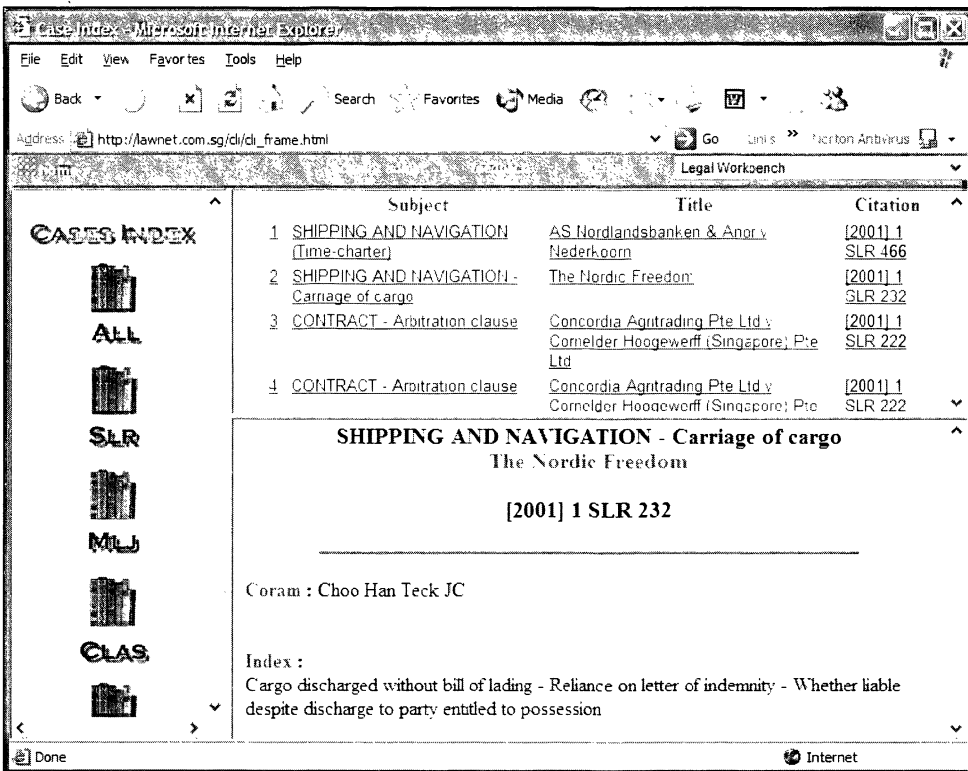


Figure 4: Cases Index

Cases Index is a simple system which makes available to users the catchwords identified by the editors of the Singapore Law Reports. Instead of having to search against the full text of the Singapore Law Reports in Legal Prospector, users can access summary information from the Cases Index system.



Technically, Cases Index is built on Lotus Notes, using a simple database which is updated by staff at the Attorney-General's Chambers. The Notes database is then replicated over to the LawNet servers at regular intervals. The index keywords are therefore re-keyed in, even though the information already exists in the Legal Prospector system.

### Results of Magistrates Appeals

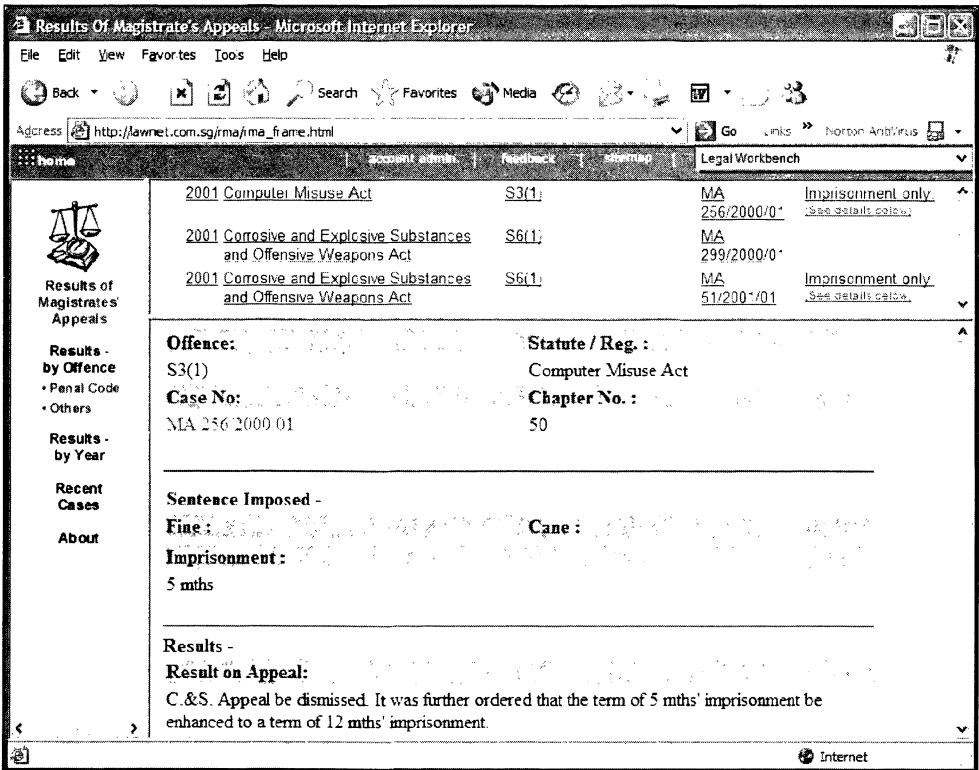


Figure 5: Results of Magistrates Appeals

Magistrates Appeals are appeals to the Chief Justice from criminal cases heard in the Subordinate Courts. Easy access to the results of such appeals is important, especially for the criminal bar. The Results of Magistrates Appeals system fills this niche. It is updated periodically by staff from the Subordinate Courts. On a technical level, the database exists as a Lotus Notes database.

## Damages for Personal Injuries

The screenshot shows a Microsoft Internet Explorer browser window with the address bar displaying [http://lawnet.com.sg/dpi/dpi\\_frame.html](http://lawnet.com.sg/dpi/dpi_frame.html). The page title is "Damages For Personal Injuries Database". On the left side, there is a navigation menu with buttons for "By Category", "By Court", "Anatomy Charts", and "About". The main content area is titled "Search Results" and contains a table with the following data:

Category	Court	Year	Case	Damages
Ankle - Crushed, Knee - Contusion	Magistrate's Court	1996	June Sim Siew Yen v Benfort Enterprise Pte Ltd	Assessed Date: 2 August 1999

Below the table, there are search filters:

- Year of Accident:** 1996
- Court:** MC
- Type of Injury:**
  - (a) Contusion of the right knee.
  - (b) Crushing injury to the right ankle resulting in swelling and an ulcer over the medial aspect.
  - (c) Scars
- Damages Awarded:**
  - Assessed Date: 2 August 1999
  - (1) Pain and suffering for lower limb injuries:
    - (a) \$13,000
    - (b) \$5,000
  - (2) Pain and suffering for scars:
    - (c) \$1,000
  - (3) Loss of earning capacity: \$5,000.

Figure 6: Damages for Personal Injuries

Civil litigation practitioners need to have access to information about the prevailing levels of compensation for various bodily injuries, separately or in combination. As a result, the Damages for Personal Injuries (“DPI”) database was created. The Subordinate Courts compiles this information, and provides a soft copy of the tables compiled and collated to the LawNet Secretariat. Staff at the Secretariat key this information into a Lotus Notes database.

## The Heritage Law Reports

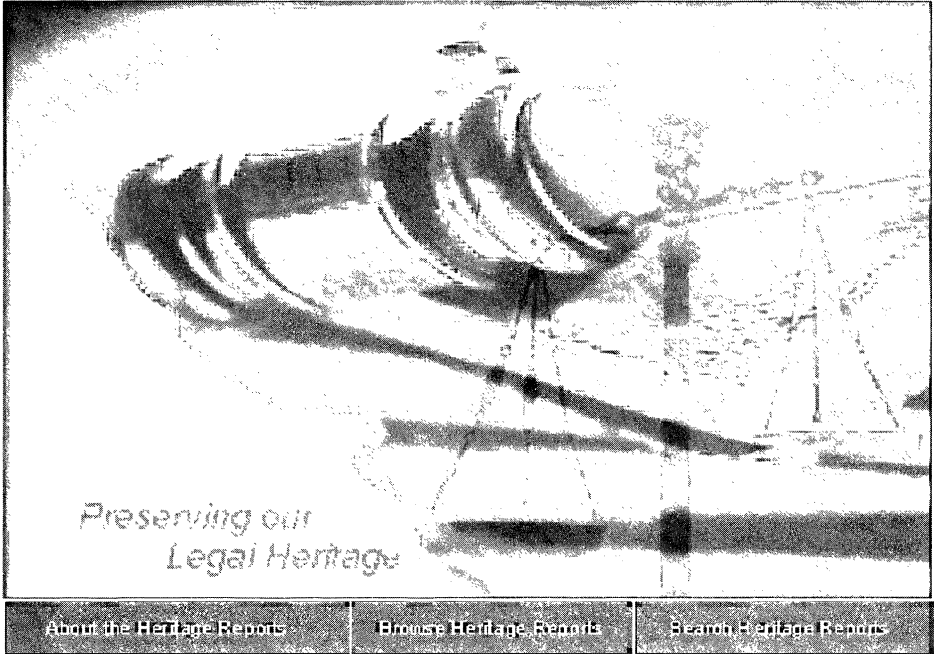

**LAWNET Heritage Law Reports**


Figure 7: The Heritage Law Reports

In 2000, when the tenth anniversary of LawNet was celebrated, the Heritage Reports were also launched. This service arose from the Attorney-General's assessment that the older law reports were at risk of being lost to the legal community because of their rarity in hard copy, as well as the deterioration of whatever hard copies existed. As a result, the Attorney-General financed the data conversion effort and arranged for the Heritage Law Reports to be placed in the LWB. The reports concerned pre-date the inception of the *Malayan Law Journal*, and date back to reports available in the nineteenth century.

On a technical level, the reports, after data conversion, were placed into the Legal Prospector backend system, as a combination of meta-data in a SQL database and raw HTML files. From the perspective of end users, a separate front end for Heritage Law Reports supplements the Legal Prospector interface. This allows for researchers whose

primary interest lies in such historical data to focus on these historical records.

The workflow requirements for this service are minimal as there are no regular updates unless new reports are found and converted for inclusion into the Reports.

### The Military Court of Appeal Cases

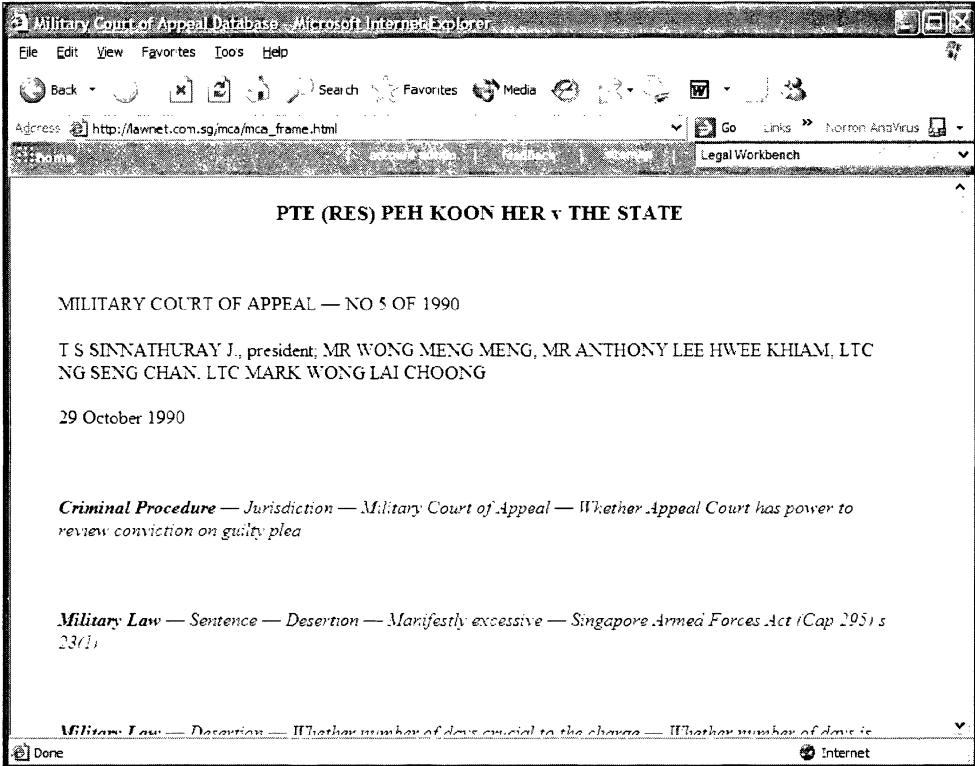


Figure 8: Military Court of Appeal Cases

Singapore has compulsory military service. Even after they are discharged from active full time military service, male Singaporeans are still subjected to regular recalls for training and skills upgrading. As a result, almost all Singaporean males are subjected to military law where their military

1 See “Does Your Company Have a Content Management Problem?” at [http://www.cmswatch.com/Features/Opinion\\_Watch/FeaturedOpinion/?feature\\_id=45](http://www.cmswatch.com/Features/Opinion_Watch/FeaturedOpinion/?feature_id=45) (visited 25 November 2001).

service is concerned. Disciplinary and other offences can be adjudicated by general courts martial, and on appeal, the Military Court of Appeal.

Through collaboration with the Director of Legal Services of the Ministry of Defence, who is the Chief Military Prosecutor, LawNet arranged for all Military Court of Appeal decisions to be available through LWB. Soft copies are received from the Ministry, and processed by the LawNet Secretariat. From a technical standpoint, the data is stored in HTML directories in IIS, and indexed with Microsoft Index Server. Workflow requirements are not significant as there are only a handful of written judgments from the Military Court of Appeal each year.

## Challenges

### WHAT ARE THE CHALLENGES?

With the benefit of hindsight and the clarity of the perfect vision that comes with it, it should be apparent that there is tremendous overlap in the case law databases within the LWB. Different services were developed by different teams of developers, at different times. Sometimes, even when developers knew of a related development for a different module within LawNet, they still opted to create different databases. There was a tendency to utilise the easiest development approaches. Typically this involved the use of Lotus Notes databases, which could be created and modified easily with very little raw programming required. Finally, the need for different people, either in different organisations, or even in the same, to perform data updating duties meant that it was easier to create distinct and separate systems for different databases.

The price for expediency is significant, and only apparent when quality problems crop up, or product improvement is considered. Some problems are as follows:

- Each case law database repeats basic meta-data about each case—apart from the duplication in effort, this represents a data accuracy risk—the same case name might be entered slightly differently in each database, and case name consistency becomes a many-time challenge instead of a one-time challenge. In addition, the work effort required to enter the same meta-data into different databases is redundant, and increases the time taken to turn around data.
- None of the case law databases is “aware” of whether or not a particular case also exists in another database. For example, the Academy Digest system has no means of knowing that a particular case has already been reported, and exists in the Legal Prospector

system. Conversely, the Unreported Judgment entry for a case will not recognise that the same case has been briefed in the Academy Digest system, or that Personal Damages information has been collated within the Damages for Personal Injuries system. As such, when viewing a record, the user cannot be referred to other related records in the other case law databases.

- It is not possible to consolidate all entries pertaining to a single case, and to display all available documents pertaining to that case in a comprehensive, coherent display.
- Workflow management for the databases is difficult to track. There is no “helicopter” overview of the status of all the databases, with workflow tracking as to where data may have been delayed in the content creation-verification-upload-to-production chain
- It is not possible to allow users to “subscribe” to obtain email notifications of new content, without either introducing additional systems which add further complexity to the system, or manually compiling such notifications.
- There is no consistent or predictable Uniform Resource Locator pattern for the documents contained in LWB. Furthermore, Lotus Notes databases are plagued with incredibly long and cumbersome URLs which are difficult to cite accurately.
- The search engine syntax varies between the MS Index Server and the Lotus Notes Search Engine. This leads to user confusion.
- It is also not possible to carry out cross-database searches, or meta-searches.
- The search engine interface, which is intended to insulate users from varying search engine syntax, also varies from submodule to submodule.

## RATIONALISING AND CONSOLIDATING DISPARATE DATABASES

The effort to consolidate all these databases is a very substantial one. This effort is further complicated by the need to cater to data conversion of all the documents already existing in the current systems. One relatively easy way to resolve this is to accept the possibility that the transition between the “old” (i.e. the current) system and the “new” system, may require a bifurcation of user access—a period of time where users are directed to different search engines for different content, in accordance with a cut-off date which can be rolled back when data conversion takes place.

In conceptualising a new and better system, one of the first “revelations” was that the entire structure of the disparate databases breeds duplication and the proliferation of data mismatch or outright errors. Each case (with a written judgment) that emanates from the courts should exist as a single and unique entry in a master database. This unique entry should contain all the necessary meta-data that identifies the case, and that is available as soon as a judgment is issued (for example the names of the parties, judgment date, *coram*, lawyers etc). This information should only need to be keyed in once for all the various submodules. Each constituent submodule should then leverage off this basic meta-data, either through the addition of further content variants to the same database entry, or through association handled and managed through a master database key or unique document ID which links the case through all the different databases.

#### A UNIFIED CASE LAW DATABASE

It is possible to visualise, at a high level of abstraction, a unified, single case law database which operates as follows:

1. When a court is about to hand down a written judgment, the personal secretary or the judge obtains a document number or a neutral citation reference number from the unified case law database. In order to do so, the unchanging fields pertaining to that judgment are entered. These may include: names of parties, *coram*, date of judgment and court case number.
2. The judge may only issue the judgment when a reference number is obtained from the system.
3. The judge issues the judgment and either uploads the Word document to the system or forwards it to a LawNet officer to work on.
4. The document first presented to users is a PDF version of the Word document, as the first variant.
5. LawNet Secretariat works on the Word document, converts it into disciplined HTML or XML format and uploads this as the second variant of the judgment.
6. The judgment is sent for summarising into digest format. When the Academy Digest entry for the judgment is available, it is uploaded as a third variant of the judgment.
7. Judgments that are selected for reporting are sent to be processed by the appointed Singapore Law Reports publisher, then annotated and typeset for hard copy publication.

8. The SGML version is uploaded when the publisher has finalised it. The PDF corresponding to that SGML of the Singapore Law Reports variant is also uploaded (or generated from the SGML variant). These are the fourth and fifth variants respectively.
9. Additional data based on each matter can be uploaded as further variants, differentiated by document type (eg a digest of damages for personal injuries information or a record of results at magistrates appeal stage).
10. It is possible that summary information about cases (eg assessment of damages, results of magistrates appeals) may be created for cases even where there is no written judgment. In these cases, the system can prompt the user for the additional meta-data fields which would not have been entered in (1) above.
11. Additional related documents (eg subsequent cases that reference a particular case) can also be added as further reference points in the record of any case in the judgments system. Additional documents may also be secondary materials (articles, textbooks etc), that reference the case in question.

The fact that the system stores all the files and all the variants against a single instance of the basic meta-data for each case does not preclude the creation of separate “views” of the data. Hence it should be possible to have an entry point into the system that only displays cases where a Singapore Law Reports variant exists. However, it is likely that with such a unified case law database, users will gravitate towards this all in one system, from which they can obtain any of the several variants of information regarding a particular case.

#### TO WHAT END?

The advantages of such a system are clear:

- Instead of multiple databases there is only a single database to manage at the backend. This reduces the proliferation of database and database systems;
- Instead of multiple search engines, each with a different search syntax, there can be a single overarching search engine, which can manage and index all the content, and return hits restricted to specific collections of documents where necessary. Conversely this means that meta-searches or cross-collection searches will be available as a matter of course.
- A single system means that there will not be multiple points of potential failure, each of which may affect the integrity of the LWB



as a whole. Instead, unscheduled downtime for a single system is easy to identify, and much less troubleshooting to identify the source of the downtime is required.

- The opportunity to rework the existing systems would allow for the introduction of new services, like email notification of new content, or new content matching search criteria.
- New content management systems can help ensure timeliness in data and proper content checking and content creation and approval workflow, with email notifications at each step of the content creation chain.
- It could be a requirement of any new system that it must allow for easy or predictable URLs, or at least short ones which can be cited with ease.
- It could also be a requirement that any new system must be XML-compliant so that it can be easily adjusted, through the use of XML stylesheets, to deliver content, formatted and abbreviated wherever appropriate, for different devices (eg PDAs, WAP phones etc).

## TECHNICAL OPTIONS—HOW TO GET TO THERE FROM HERE

- Content bottlenecks with a webmaster, IT department, or some other gatekeeper.
- Site visitors have difficulty finding what they need.
- Content contributors have difficulty finding what they need.
- Some content is inaccurate / outdated / redundant / unauthorized.
- The website exhibits inconsistent design and navigation schemes.
- The home page does not provide a full, up-to-date portal into the rest of the site.
- Contributors occasionally overwrite content / files accidentally.
- Web managers need to "roll back" the site to a previous version -- perhaps for legal or regulatory reasons -- but cannot.
- Content contributors are unable to pre-publish content to appear at a specified later date or time.
- Website managers cannot associate the company's products and services to articles or news on the site (or vice-versa).
- Content has feet of clay: web managers cannot easily reuse / share / distribute / import it.
- An inability to protect or control access to content keeps good material offline.
- Marketing and product managers cannot customize content for customers, partners, and other important visitors.
- Internal company staff is not invested in web communications.
- Company staff lament, "Our website is not as good as we are."

What? MFDiagnosix Reassess

In exploring the collapsing of the various case law databases, many technological options were explored. A vendor initially proposed that we simply embrace the use of the iPlanet Application Server as a framework for what can only be presumed to be the wholesale custom coding of the system. Eventually, we came to hear of the current buzzword, "content management systems".

Figure 9: Self diagnosis checklists are available on the web

One website succinctly described the problems we encounter in managing the content that is put on LawNet in general, and more specifically within LWB. Another gave us a nice self help checklist which quickly assessed us as having “a bona fide content management problem”.<sup>1</sup> Gartner, Forrester and Ovum all have copious reports discussing and analysing content management systems, and how they can revolutionise the management of large websites.

One of our requirements was that the content management system must have local representation within Singapore. On the basis of our discussions with vendors, two products have risen to the top of our shortlist. Both claim to be managing extremely high profile websites—either of newspapers, top tier banks, or other top companies. In addition, open source equivalents are also being explored. An open source content management initiative, called Midgard, working with scripting languages like PHP, has also been trialled.

Whilst we are aware of the incredible success of AustLII, our study of the rich documentation seems to indicate that AustLII has different priorities to ours. Whilst AustLII has a remarkable search engine, SINO, our needs seem to lie more with the control and management of disparate but related content. In addition, the philosophy of AustLII is “free to air” access and whilst the situation may change, our current operational paradigm is still focused on the provision of a commercial, chargeable service.

Vendors have plied us with brochures, glitzy PowerPoints and demonstrations. We are still evaluating the various options, but the following factors and questions should be asked by anyone looking to re-engineer case law databases through the use of a content management system:

- Should content be static or dynamically generated from a database?
- How flexible are the workflow approval chains required?
- Does the solution work best with a two tier or a three tier architecture?
- How easy is it to create new templates?
- How configurable are the security and “role” settings?
- Does the solution support XML and XSL technologies? Do you need it to?
- What is the user load which the delivery servers must be able to sustain?
- Will multiple content creators be needed to amend the same content?

- To what extent is rollback available on the system? Can the entire site be subjected to rollback? Specific areas or collections? Or only individual assets?
- Thesauri and taxonomies are important considerations in any disciplined content management paradigm.
- Does the solution make use of any proprietary solution or content, or does it make use of open standards as far as possible?
- How would data migration away from the system be affected if, at some point in time in the future, the solution has to be abandoned?