

Exhibit K

(Selected Pages of Deposition of Timothy Arnold-
Moore, with Deposition Exhibits 1 and 8)

to

TimeBase's Memorandum in Support of Its Motion
for Summary Judgment of No Invalidity

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF MINNESOTA

-----)
TIMEBASE PTY LTD.,)
Plaintiff,) File Number
v.) 07-CV-1687
THE THOMSON CORPORATION, WEST)
PUBLISHING CORPORATION, AND WEST)
SERVICES, INC.,)
Defendants.)
-----)

Washington, D.C.

Tuesday, June 22, 2010

Videotaped Deposition of:

TIMOTHY J. ARNOLD-MOORE, Ph.D.,
a witness called in the above-entitled action,
before BESS A. AVERY, RMR, CLR, a notary public in
and for the District of Columbia, taken at Morgan,
Lewis, & Bockius, 1111 Pennsylvania Avenue,
Northwest, Washington, D.C., commencing at
9:35 a.m., when were present:

122

1 THE WITNESS: Thank you.
 2 MR. HOSTENY: Let's take a break.
 3 THE VIDEOGRAPHER: Are we concluded?
 4 MR. LITSEY: No.
 5 THE VIDEOGRAPHER: We are going off the
 6 record at 12:19 p.m.
 7 (A luncheon recess was taken.)
 8 THE VIDEOGRAPHER: This marks the
 9 beginning of Videotape Number 4 in the deposition of
 10 Dr. Timothy J. Arnold-Moore. The time is 13:10
 11 hours.
 12 EXAMINATION BY COUNSEL FOR THE PLAINTIFF:
 13 BY MR. HOSTENY:
 14 **Q Dr. Arnold-Moore, am I correct that when**
 15 **you use the word "consolidation" you are referring**
 16 **to an act with one or more amendments applied to it?**
 17 A That's correct, yes.
 18 **Q Okay. And that would be consistent, I**
 19 **think, with if you take a look at your Exhibit**
 20 **Number 3, page Roman numeral 5.**
 21 A Yes.
 22 **Q Okay. So the consolidation will represent**

123

1 **another version of an act with all of the**
 2 **intervening amendments applied?**
 3 A That's correct, yes.
 4 **Q Do you sometimes call the change a delta?**
 5 A I certainly in some of my papers talk
 6 about the difference between two documents as being
 7 a delta, yes.
 8 **Q Okay. Would the delta represent the**
 9 **actual change of text whether it's an addition or a**
 10 **deletion?**
 11 A That's correct, yes.
 12 **Q So a delta would be, for example, if it**
 13 **were a deletion, it would be some lined out text or**
 14 **some text to be removed?**
 15 A That's usually how it's represented, yes.
 16 **Q And if it were an addition then you would**
 17 **have some sort of insertion to be made to what the**
 18 **section was?**
 19 A That's correct, yes.
 20 **Q When you say "act" what do you mean?**
 21 A An act of parliament, typically. In some
 22 of my papers when I use the term "act" I'm using it

124

1 in a broader sense of any legislative document. As
 2 I explained before, in Australian all the
 3 distinction between regulations and primary
 4 legislation or acts of parliament or acts of the
 5 legislature is not as stark as it is here in the
 6 United States, the separation of powers is not as
 7 clear-cut. So what I say about acts typically
 8 applies to regulations in those jurisdictions. But
 9 typically by act I mean an act of parliament.
 10 **Q Okay. Something passed by a legislative**
 11 **body?**
 12 A Something passed by a legislative body,
 13 exactly.
 14 **Q Okay. And you used a term in your direct**
 15 **examination to refer to the equivalent of**
 16 **regulations. I don't recall what it was.**
 17 A Well, there are a number of terms:
 18 Subsidiary legislation or subordinate legislation,
 19 statutory rules is another, another phrase I think
 20 that's used in some of the papers.
 21 **Q Okay. In those instances would those also**
 22 **be things enacted or passed by a legislative body?**

125

1 MR. LITSEY: Objection, lacks foundation.
 2 A They are typically not actually passed by
 3 a legislative body but made under a power granted by
 4 an act so they are made by an executive branch,
 5 ministry or department, and under the authority of
 6 one or more acts of parliament.
 7 BY MR. HOSTENY:
 8 **Q Okay. If you would take a look at**
 9 **Exhibit 1 for a moment, please. And I noted down at**
 10 **the bottom right, well, there's a URL reference down**
 11 **at the bottom right. Do you see that?**
 12 A Yes.
 13 **Q Do you also see the date down at the**
 14 **bottom right, June 22, 2010?**
 15 A Yes, I do.
 16 **Q Do you know where this document comes**
 17 **from?**
 18 A I believe it's from a web page.
 19 **Q Do you know who printed it?**
 20 A No, I don't actually.
 21 **Q Do you know when it was printed?**
 22 A I presume, from that date, on June 22,

126

1 2010, but I don't know.

2 **Q Okay. You mentioned that this was a,**

3 **reflected a presentation that you participated in in**

4 **September of 1996 as reflected at the third page of**

5 **the document?**

6 A Yes, that's correct.

7 **Q Okay. Were you able to find a paper copy**

8 **of the presentation as given on September 25, 1996?**

9 A No, I was not.

10 **Q You mentioned that there was a booth at**

11 **this same event?**

12 A That's correct, yes.

13 **Q Were you able to find any record of the**

14 **visitors to the booth?**

15 A I didn't maintain a record of visitors to

16 the booth, no.

17 **Q Do you know if there is any record of**

18 **persons to whom -- let me back up a bit.**

19 **I think you mentioned a one-page handout?**

20 A My memory is that there was a one-page

21 handout.

22 **Q Were you able to find that one-page**

127

1 **handout?**

2 A No, I don't believe we kept a copy of that

3 handout.

4 **Q And do you recall if there was any record**

5 **of persons to whom the handout was provided?**

6 A It was sitting on the table, so anybody

7 could have taken a copy without us taking a record

8 of who had.

9 **Q Okay.**

10 A I could provide -- I could provide a

11 partial list of people who I know visited the booth,

12 but it would be from my own memory. I don't believe

13 there's any written record. There may be a written

14 record of attendees to the conference itself that

15 Allette Systems had that ran the conference.

16 Allette Systems is a company based in Sidney that

17 ran this particular conference.

18 **Q Okay. I want to ask you a few questions**

19 **about EnAct. EnAct dealt with legislation.**

20 **Correct?**

21 A That's correct, yes.

22 **Q And in particular the legislation of**

128

1 **Tasmania?**

2 A As it was originally created, yes.

3 **Q And a single jurisdiction, that**

4 **jurisdiction being Tasmania?**

5 A That's correct, yes.

6 **Q And you could, as part of EnAct, explore**

7 **how legislation varied over time?**

8 A That's correct, yes.

9 **Q Okay. Give me a minute here while I**

10 **thumb.**

11 A Sure.

12 **Q If I understand -- you want to flip to**

13 **Exhibit 8.**

14 **Exhibit 8 is a presentation given at some**

15 **time in 1999 concerning EnAct?**

16 A That's correct, yes.

17 **Q And EnAct came online in 1998 or 1997?**

18 A I believe --

19 MR. ROTH: Objection.

20 A I believe it was 1997.

21 BY MR. HOSTENY:

22 **Q Okay. Let me direct your attention to --**

129

1 A Actually, by online you mean in production

2 in the Tasmanian Government, or do you mean the

3 website was available to the public?

4 **Q Well, let's split it up. According to --**

5 **take a look, if you would, at Exhibit Number 15, the**

6 **Legislative System Project Newsletter.**

7 A Yes.

8 **Q It says that the implementation of the**

9 **EnAct legislation system went, "into the office of**

10 **Parliamentary Counsel on 1 December 1997."**

11 **You see the upper left paragraph under the**

12 **production?**

13 A Yes, mm-hmm.

14 **Q So was it available as of December 1,**

15 **1997?**

16 A It went into production on the first of

17 December 1997 in the Office of Parliamentary

18 Counsel. They were using the system in system

19 testing well before that.

20 **Q Okay. How long before, do you know?**

21 A At least six months and possibly up to a

22 year.

130

1 **Q** Take a look back at Exhibit Number 8, if
 2 you would, the second page.
 3 A Second page, yes.
 4 **Q** Look at the abstract there.
 5 A Mm-hmm.
 6 **Q** See where it says, I'm not quoting, but it
 7 says the point-in-time capability allows users to
 8 search and browse, et cetera, as it was at any time
 9 since 1 February 1997?
 10 A Mm-hmm.
 11 **Q** Okay. So if I looked at it at 1 February,
 12 1997, I would not be able to access any legislation
 13 prior to that date, would I?
 14 MR. LITSEY: Objection, lack of
 15 foundation.
 16 A The -- that particular date was chosen as
 17 a date before which Tasmania would not prepare all
 18 the versions of the document, but, in fact, if you
 19 search for legislation before that date, there was
 20 some legislation that was actually in force, at
 21 earlier dates of the complete collection.
 22 BY MR. HOSTENY:

131

1 **Q** When was EnAct online, available to any
 2 user who wished to consult it?
 3 MR. ROTH: Objection.
 4 MR. LITSEY: Objection, lacks foundation,
 5 vague.
 6 BY MR. HOSTENY:
 7 **Q** Let's back up. Was it ever online?
 8 MR. ROTH: Object.
 9 A There was a website that was built from
 10 the EnAct system that was made available in 1998
 11 that was available -- which made Tasmanian
 12 legislation and the contents of the EnAct database
 13 available to the public in 1998, but it depends what
 14 you mean by "online."
 15 BY MR. HOSTENY:
 16 **Q** I mean available to anyone who wished to
 17 determine the content of Tasmanian legislation
 18 electronically?
 19 A As a website it was available in 1998.
 20 **Q** Okay. Page -- look at Figure 1, if you
 21 would. It's several pages into Exhibit number 8.
 22 A Mm-hmm.

132

1 **Q** That's an example of a, that Figure 1 is
 2 an example of a paper Markup of -- let me back up
 3 That Figure 1 is an example of a paper
 4 Markup to achieve a consolidation?
 5 A That's correct, yes.
 6 **Q** Okay.
 7 A In fact, I think that particular image was
 8 actually scanned from one of the paper documents
 9 that was being used in the Tasmanian Drafting Office
 10 before the EnAct system went live.
 11 **Q** Okay. In any of the exhibits that you've
 12 seen today is there any discussion of examining or
 13 -- let me rephrase that.
 14 **Is there any discussion of studying**
 15 **legislation in variations other than time, if you**
 16 **follow what I'm saying?**
 17 MR. LITSEY: Object as vague and lacks
 18 foundation.
 19 A Well, I think I talk about other aspects
 20 of legislation in some of those papers; for
 21 instance, the regular structure, the fact that it
 22 involves cross-reference. I made quite a tight

133

1 reference of cross-reference links. So time is one
 2 of the aspects that's relevant, but in some of the
 3 papers I discuss other aspects of legislation that
 4 make it different from other document types.
 5 BY MR. HOSTENY:
 6 **Q** Okay. Let me find that. I want to ask
 7 you about that figure, while I'm at it. Give me a
 8 moment. What I'm looking for is the document that
 9 had the figures in it that you were describing to
 10 Mr. Litsey.
 11 A The screenshots?
 12 **Q** Yes, that's it.
 13 A That would be Exhibit 8.
 14 **Q** Is that 8? Okay. Let's go back to that
 15 one. I think you are probably right.
 16 A Is it the AustLII paper? No, it was the
 17 digital libraries paper. 7, my mistake.
 18 **Q** Seven. Okay. By the way, what is Themis?
 19 A The EnAct system was originally called
 20 Themis while we were working on it as a project.
 21 There were some trademark issues with using that
 22 name so the name was switched to EnAct.

142

1 A Of Themis, yes. It then became EnAct.
 2 **Q When I say it's new, did you ever become**
 3 **aware of any system that was available prior to**
 4 **Themis that did that?**
 5 MR. LITSEY: Object as lacking foundation.
 6 A The prior search which is a research
 7 system, I think from Carnegie Mellon, I couldn't be
 8 sure, does some similar things. It doesn't
 9 necessarily -- it's not focused on SGML but it does
 10 break up hypertext documents into nodes which
 11 roughly correspond with elements in an SGML
 12 document.
 13 So there was some -- there were some
 14 systems that did actually store documents in pieces
 15 and had hypertext links between them before EnAct.
 16 So would I say it was new in the sense that nobody
 17 had ever done it before, probably exactly the way we
 18 did it, no, but certainly the concept of breaking a
 19 document up to store it in a hypertext system, that
 20 wasn't novel.
 21 BY MR. HOSTENY:
 22 **Q Okay. How many years did you and your**

143

1 **colleague spend working on Themis and EnAct?**
 2 A I believe the project started in, it was
 3 either late '94 or early '95. We started -- we were
 4 awarded the tender. The final system was delivered
 5 in December, I think, first of December '97, one of
 6 the exhibits said that was when the system went
 7 live. Complete versions of it had been delivered
 8 before that but there was system testing and fixes
 9 and so on before that, so probably a period of two
 10 years, roughly.
 11 **Q Was it your belief based on your work on**
 12 **Themis to the effect that there was not an available**
 13 **system that could accomplish the same thing that**
 14 **Themis could accomplish?**
 15 A There certainly wasn't an off-the-shelf
 16 solution that could do all the things that Themis or
 17 EnAct could do.
 18 **Q Yeah. Okay. Now, you mentioned at page**
 19 **179 the use of delta operations.**
 20 A Yeah.
 21 **Q And chained delta operations.**
 22 A Mm-hmm.

144

1 **Q What's the difference between a delta and**
 2 **a chained delta?**
 3 A Well, version control systems at the time
 4 use the concept of delta to describe an encoding of
 5 a description of the changes between two versions of
 6 a document. So a delta was, effectively, something
 7 that you could apply to one document to get the next
 8 version.
 9 And the idea of chain deltas is that,
 10 well, if there's multiple versions between the
 11 version you are interested in and the version that
 12 you had, you would apply a chain of those deltas to
 13 get the final version you're interested in. Or, in
 14 fact, typically, in versions for old systems, it was
 15 the other way around, you kept the most recent
 16 version and chained back to the old versions.
 17 **Q Why are you mentioning deltas and chained**
 18 **deltas in here?**
 19 A Because that was the language that was
 20 being used by the Version Control Systems at the
 21 time, and there were some parallels between Version
 22 Control Systems and what we were trying to achieve

145

1 with EnAct.
 2 **Q Did you use what you would call an**
 3 **inverted file index in Themis?**
 4 A Yes, the Structured Information Manager,
 5 as it was called then, or TeraText database at the
 6 moment, it uses an inverted file index to index all
 7 the words that appear in a record that's indexed.
 8 **Q Did Themis use an inverted file index?**
 9 A It did. It made use of that.
 10 **Q Does EnAct still use an invert --**
 11 A It still does, yes.
 12 **Q Is an inverted file index an index of**
 13 **individual words appearing in a document?**
 14 A That's correct, yes.
 15 **Q And SIM, that's the Structured Information**
 16 **Manager?**
 17 A Yes.
 18 **Q Did that become TeraText?**
 19 A That's correct, yes.
 20 **Q And TeraText was a product of InQuirion,**
 21 **if I pronounced that correctly?**
 22 A Yeah, well, the Structured Information

[Mirrored from: http://www.gca.org/conf/asia_pac/program.htm]

SGML Asia-Pacific Final Program

Monday, 23 September: Pre-Conference Activities

9:00-16:30

OmniMark Tutorial and User's Group Meeting

17:00-19:00

Early Registrant Cocktail Party

For those arriving the day before who wish to meet fellow attendees and complete registration prior to the start of the Conference.

Tuesday, 24 September	Wednesday, 25 September	Thursday, 26 September
------------------------------	--------------------------------	-------------------------------

Tuesday, 24 September

8:00-12:00

Conference Registration

9:00-12:00

Pre-Conference Tutorials

12:15-14:45

Conference Welcome and Opening Keynote Luncheon

Inventor's Keynote

Dr. Charles Goldfarb, *President, Information Management Consultants and Inventor of SGML and HyTime, USA*

In a talk that has become a tradition at the SGML conferences in Europe and North America, Dr. Goldfarb, the inventor of SGML and HyTime, discusses current challenges and opportunities facing the SGML community. The state of SGML-related standards activities will also be covered.

SGML in a Modern World

John McFadden, *President, Exoterica and Chairman of the SGML Open Board of Directors, Canada*

The advent of CD-ROM and WWW technology has driven the need to develop new ways to collect, deliver, and browse information. The practice of SGML is rapidly evolving to keep pace with the demands of users. John McFadden will discuss the emerging importance of SGML as a component of a larger information technology framework in a modern client/server environment.

14:45-15:00

Break

15:00-17:00

Concurrent Session A: SGML for Multilingual Publishing



Chair: Robin Masson, Uniscope, Japan

Multilingual Document Interchange Using SDATA Entity and Others

Yasuhiro Okui, Nihon Unitec, Japan

Multilingual documents include composite character sets whose coding are standardized or non-standardized. Several schemes for treatment of those non-standardized characters and character sets are discussed from the SGML document interchange's point of view. Some guidelines for use of coded character sets, SDATA entities and glyph identifiers are proposed.

Multi-language Conversion using SGML: A Case-Study

Bernd Nordhausen, Passage Systems, USA

The conversion of documents in different languages into SGML presents additional challenges beyond those encountered in a conversion effort of single language. This is especially true when one enters the realm documents encoded in different characters sets including multi-byte sets. In this paper, we discuss an effort of converting about 200 books totalling 9,000 pages into 15 languages including 11 European, and four Asian (Korean, Japanese, simplified and traditional Chinese). This presentation describes the tools used in this conversion effort and reviews the practical issues involved in multi-language SGML conversion.

Asian Language Character Set Processing Issues

Rick Jelliffe, Allette Systems, Australia

With technology developed for processing roman character sets, there is often predictable lag before counterparts are released to process Asian character sets. This presentation will look at working with SGML to process Asian language character sets using non-SGML applications. This presentation will discuss practical examples of east Asian character language formatting and examine standards issues for the future of Asian language processing.

How to Implement SGML for an Operation and Maintenance Manual - the Ericsson Example

Ulla Berg, Ericsson, Japan

Ericsson Radio Systems AB is delivering a Mobile Telephone System to Japan, the CMS 30 System. This presentation describes how SGML was implemented for the 40 volumes of Operation and Maintenance documentation. The Operation and Maintenance Manual is produced in English and then translated to Japanese. The presentation will also describe how to implement SGML for the documentation in Japanese.

15:00-17:00

Concurrent Session B: SGML Document Management

Chair: Nick Carr, Allette Systems, Australia

Integrating an SGML Solution into a Workflow Environment

Graham Tritt, Swiss Federal Office of Information Technology and Systems, Switzerland

The administration of legal documentation is ideal for a solution based on the technologies of databases, SGML and workflow management. The topic of this presentation is how these different technologies interact. It will discuss what exactly can be usefully stored and searched with the relevant type of database, what format conversions can be automated, and how the management of the data can be integrated into the bigger framework of a work-flow system.

Management of SGML Data at QANTAS Engineering & Maintenance

Bill Donoghoe, Computervision, Australia

This presentation will look at how Computervision has implemented an advanced data management system for the QANTAS Engineering Training School. The core component of this system is the Optegra Enterprise Data Management software which provides controlled access & updating of all types of data, including SGML. Major topics to be covered in this presentation include:

- the data model which provides single point data maintenance and the ability to construct

- training courses as required through different views of the database;
- migration of the existing file-based SGML repository into an object-oriented, client-server environment; and
- the use of a phased approach to the implementation of this system.

Building a Document Management System: The issue of Interoperability

Brendan Hills and Anne-Marie Vercoustre, *CSIRO Division of Information Technology, Australia*

Document standards such as SGML provide a format which allow work on the same document with different programs. Systems such as OLE and OpenDoc or IBM's "BluePrint" are very useful in providing a mechanism for software to interoperate when that software is designed with interoperability in mind. What we have found however, is that many of the tools in which we are interested have not been designed to provide this software level of interoperability. It is this finding which has lead us to develop a system for third party application integration. It will be this development and the issue of interoperability that will be the subject of this presentation.

17:00-19:00

Exhibits/RECEPTION

Wednesday, 25 September

9:00-10:15

General session: SGML Databases

Chair: Brian Travis, *Information Architects, USA*

Legislation in its Natural State

John Cook, *Ferntree Computer Corporation, Australia*; **Tim Arnold-Moore**, *Royal Melbourne Institute of Technology, Australia*; and **Sharyn Paice**, *Tasmanian Department of Premier and Cabinet, Australia*

This presentation will examine the business imperatives of the legislative drafting process, the end-user and technical requirements and the decision making that led to the adoption of SGML and the realization of benefits. It will look at the technical architecture of the solution for drafting, information retrieval, legislative amendments and consolidations and the document database. And it will discuss the technical hurdles involved in integrating the various modules, the limitations of the components and the development of automatic consolidations.

Hybrid Document Databases - Integration Solutions for SGML and Relational Database Technology

Eric Skinner, *Exoterica Corporation, Canada*

Large-scale SGML implementations require integration with database systems in order to perform information management tasks. A practical RDBMS/SGML hybrid model, Hybrid Distributed Database (HDDDB) will be discussed. HDDDB enables your publications and RDBMS information to be part of a single information model, reducing redundancy and enhancing reusability. A multi-vendor working application will be demonstrated.

10:15-10:30

BREAK

10:30-12:30

Concurrent Session A: CALS

Chair: Nick Carr, *Allette Systems, Australia*

Document Management in the Defense Environment

Ken Holman, *Microstar, Canada*

Microstar has been long recognized as an innovator in DTD design, and have discussed their work with the Canadian Department of Defense at previous conferences. In this presentation, Ken Holman will explain the progress of a leading military SGML implementation with particular reference to document management.

SGML in an Integrated Logistics Environment**Mike Bowden, *GEC Marconi Systems, Australia***

This presentation will explore a customer-focused perspective on the use of Interactive Electronic Technical Manuals (IETMs) to provide information to engineers "in the field." There are numerous situations related to through life support that demand fast, flexible access to information to solve problems. Data such as parts lists, prices, manufacturer's part specifications, etc. are essential for providing comprehensive integrated logistics support. Discussion will center on the business benefits of efficient information delivery through the use of SGML.

Caterpillar Inc's New Authoring System**PG Bartlett, *ArborText, USA***

Caterpillar, Inc. has developed a new document information system that emphasizes the reusability of Information Elements (shared objects) in multiple documents, the automatic compilation of objects into a document, and the reusability of documents on multiple media. Based on ISO and military standards, the new information system will improve accuracy, consistency, efficiency, timeliness and costs. This paper describes the issues that led to the system's design, pitfalls in its implementation and operation, and details the anticipated benefits.

CALS as a Foundation for Advanced Information Delivery**Major Dale Bradshaw, *Australian Army***

For the last three years, the Australian Army has been involved with converting hardcopy training manuals to SGML using a customized version of the CALS DTD. As the library continues to build, the Army sees the main challenge as maximizing the value of the SGML investment. This presentation describes how various tools and services such as multimedia, on-line delivery and object management enables both domestic and international users with varying degrees of expertise to utilize the SGML data in a variety of ways.

10:30-12:30**Concurrent Session B: SGML Issues for Complex Data Structure****Chair: Julia Yeong, *International Publishers Distributors, Singapore*****SGML Database Technology****Andrew Ogbourne, *Butterworths, Australia***

One of the most significant issues facing implementors is how to manage their SGML data. Initially, the most likely solution was to simply store instances as files and this is still the preferred option for many installations. However, it is becoming increasingly attractive and feasible to manage SGML data in some form of database management system. This presentation will discuss the advantages and disadvantages of making the leap to a database for SGML and the issue of relational versus object-oriented database technology. It will also provide advice about evaluating and selecting SGML database solutions.

Legal Publishing**Seet Chern Hway, *Information Technology Institute, Singapore***

Abstract unavailable

Large-Scale Government Statistical Publishing**Knut Vidar Hoholm, *Open Systems Consultants, Norway***

This presentation will discuss an SGML implementation at Statistics Norway. The purpose of the project is to automate the process of editing, revising and publishing the 1996 Norwegian Statistical Yearbook. Focus will be on the user requirements DTD modeling and the choice of

tools. The benefits of a better quality of published data, more recent information and reduced costs will also be highlighted.

Guidelines for Academic Journals On-line

David Green and Terry Bossomair, *Charles Sturt University, Australia*

It is ironic that the one of the user communities that has heavily contributed to the creation and popularization of the World Wide Web, academic research, is also one that has the most difficulty effectively publishing on the Web. Research papers containing complex data such as tables, equations and formulas can be very difficult to get into HTML. This presentation will look at the issues associated with academic publishing on the Web.

12:30-13:30

LUNCH

13:45-15:15

General Session: SGML Conversion Issues

Chair: Robin Masson, *Uniscope, Japan*

Legacy Data And Structuring For Future Use

Raymond H. Stachowiak, *XyVision, USA*

For decision makers considering implementing an SGML-based Document Management solution, one of the most difficult obstacles to overcome is the conversion of existing data into the optimal format for efficient reuse. Citing experiences from various companies in military, government, and other industries, Stachowiak will discuss the issues of preparing existing data for use in a document management system, with emphasis on specific problems facing Asia Pacific companies.

Cost Justifying SGML

Norma Haakonstad, *ArborText, USA*

When making a business case for SGML, one of the key arguments is justifying the cost for the transition to SGML. Whether your objective is to support multiple outputs or to re-engineer your information production processes, this presentation is designed to help you justify the cost of implementing SGML. This presentation will cover the measurable benefits in detail, the unmeasurable benefits of SGML, and provide suggestions for preparing your argument.

Economics of SGML Conversion

Dale Waldt, *Research Institute of America, USA*

One of the biggest cost areas in an SGML migration is the conversion of legacy data. In spite of it's importance, it is often underestimated. Proper planning and estimating will pay off in the long run and allow a smooth transition to the new SGML-based system. Some ideas and techniques for SGML conversion will be discussed.

SGML Conversion and Document Management at Texas Instruments Semiconductor Using Pinnacles PCIS 1.1

Jeff Barton Taylor, *Texas Instruments, USA*

This paper presents a practical, working example of an SGML application which delivers printed paper, CD-ROM, facsimile, and WWW (Internet) copies of product technical information (datasheets) from a centrally managed repository via automated and semi-automated processes driven by a database. Like many database publishing systems, the system at TI Semiconductor is evolutionary. The paper presents the engineering outline of the system, including description of the source formats and the output types. It also presents the timeline and major milestones for the system development and implementation, as well as financial aspects such as equipment acquisition, contract programming and system integration costs, SGML consulting costs, and staffing costs. This paper includes information on the Pinnacles PCIS standard for information interchange.

15:15-15:30

Break

15:30-17:00

Round Table Discussions

Adhoc Roundtables for those who wish to have in-depth discussions on specific interest area.

Topics will include:

- SGML Design Techniques
- DTD Development
- Applications Development
- Conversion Issues
- Other Topics Welcome

15:30-19:00

Exhibit

17:00-19:00

Reception

Thursday, 26 September

9:00-10:15

General Session: Standards Issues

Chair: Robin Masson, Uniscope

Basaset and Charset

Tony Graham, ATLAS Consulting, USA

What do you do when you want to use a character set other than the SGML default? How can you use different characters for tags? When fed to the parser, what comes out? Using a simple example of an exotic character set with only six characters, this presentation explains BASESET and CHARSET and their relationship to the parser and processing applications. This presentation will also discuss real life applications of this example.

Graduating from File-based Document Assembly to Info-Based Document "Construction"

David Sklar, Electronic Book Technologies, USA

SGML's entity features support the assembly of a document from a collection of modules that must be precisely enumerated and absorbed in toto. Breaking out of these limitations enables publishers to graduate from mere assembly to highly sophisticated "construction," in which novel SGML documents are auto-generated via combinations of queries (in SGML infobases and/or relational databases) and SGML transformations. This presentation discusses the limitations of assembly via entities and an introduction to the concept of document construction (DC). A description of current proprietary implementations of DC engines is given as well as a proposal for the role of standards like HyTime and DSSSL in the continuing evolution of DC technology.

10:15-10:30

BREAK

10:30-12:15

General Session: SGML and HTML

Chair: Nick Carr, Allette Systems, Australia

How to Use SGML to Enrich HTML**Chris Ziener**, *Graphic Communications Association*

HTML should not be seen as the be all and end all to communication over the Web. SGML can provide a higher level of sophistication to your Web site, but it also adds a great deal of complexity. It is not always necessary for companies to use SGML, but how does a company know when the benefits outweigh the costs? The pros and cons of both HTML and SGML will be covered and a hybrid solution will be proposed.

How to get your documents on the WWW**Lani Hajagos**, *Adobe, USA*

This paper describes a project to convert a set of documents which use a derivative of the DocBook DTD into HTML. The specific conversions issues are discussed, and the implementation strategy described.

Industrial Strength Web Publishing**Bruno Pisano**, *Allette Systems, Australia*

While there are a myriad of applications that enable the fast and friendly creation of small Web sites. Creating and managing large, complex, frequently updated Web sites demands more robust, automated solutions. This presentation will outline the technical challenges of programming, processing and managing a 50,000 page Web site. It will discuss the automation of the links, advanced navigation techniques, and the issue of working with mixed types of data, from fielded information to narrative text.

12:30-14:30**Lunch, Final Keynote and Conference Close****Web Publishing, SGML's Role in Practice****Eric van Herwijnen**, *Author of Practical SGML and President of Nice Technologies, Switzerland*

The World Wide Web has taken the world by storm. Although it has increased awareness about SGML, it also raises some fundamental questions, such as "Does the Web make SGML obsolete?", "How to deliver SGML over the Web," and "How to get more flexible formatting on the Web." This presentation will address these questions in the order of their general importance.

15:00-18:00**Facility Tour and Reception**

Hosted by the Australian Army Training Corp

Friday, 27 September: Post-Conference Activities**9:00-16:30****Post-Conference Tutorials**[Return to SGML Asia-Pacific Index](#)

Registration Information



[Return to the GCA Home Page](#)



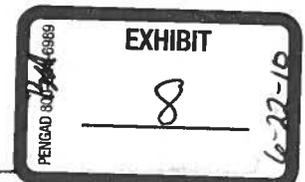
Law via the Internet '99

- 27 -

Connected to the Law: Tasmanian legislation using EnAct

Timothy Arnold-Moore,
EnAct Technical Director,
RMIT Multimedia Database Systems
tja@mds.rmit.edu.au

Jane Clemes,
ISB, Department of Premier and Cabinet,
Tasmania
J.Clemes@dpac.tas.gov.au



THOM00210045

Connected to the Law: Tasmanian legislation using EnAct

<http://www.thelaw.tas.gov.au/>

Timothy Arnold-Moore, EnAct Technical Director, RMIT Multimedia Database Systems

tja@mds.rmit.edu.au

<http://www.simdb.com>

Jane Clemes, ISB, Department of Premier and Cabinet, Tasmania

Matthew Tadd, Whole of Government Projects Unit, Tasmania

J.Clemes@dpac.tas.gov.au

<http://www.dpac.tas.gov.au>

Abstract:

EnAct is a legislation drafting, management and delivery system that has been built to enable the Tasmanian Government to provide improved legislation information services to the community. EnAct provides the community with a facility that enables cost effective public access to reliable, up-to-date, searchable consolidated Tasmanian legislation. The 'point-in-time' capability allows users to search and browse the consolidated database as it was at any time since 1 February 1997. Tasmania achieved these goals by automating much of the legislative drafting and consolidation process. This paper discusses why the Tasmanian government implemented EnAct, the concepts behind EnAct, and the technology that makes all of this possible.

1. Why implement EnAct?

1.1. How legislation used to be delivered

Before EnAct was initiated in Tasmania, the management of legislation in Tasmania was reaching a crisis point. Tasmania is a small jurisdiction with a small legal community. So it provided a limited market for legislation in any form. None of the commercial legal publishers, who maintained services for the larger Australian jurisdictions, could justify the cost of maintaining collections of Tasmanian legislation. In a democracy, the responsibility for providing access to legislation should fall on the shoulders on government, but often private enterprise takes on that responsibility. In Tasmania that responsibility rested solely on the Tasmanian government, namely the Printing Authority of Tasmania (PAT, the official government printer of legislation) and the Office of Parliamentary Counsel (OPC, who provide legal advice and drafting services to the Parliament).

Tapper states that 'a provision which is in force, and purports to state the law, may be subjected to numerous amendments'.¹ Each of these amendments describes changes

¹ Tapper (1970) "Computers and legislation" 23 *Alabama Law Review* 1-42.

to the text of the Act as it was originally enacted.² Periodically, the OPC compiled and the PAT published consolidations of these amendments or reprints of the substantive Acts, that is, they printed a version of the Act as it was at a particular time with all of the textual changes applied to it. The last full consolidation of the whole Statute Book took place in 1959. In the mid 70's, a program of rolling reprints was instigated, but a lack of financial resources and an increased demand on OPC for drafting new legislation meant that the program was way behind its targets. This caused considerable concern among the legal community in Tasmania.

As an example, immediately before the implementation of EnAct in Tasmania, the *Racing and Gaming Act 1952* was last consolidated in 1978. Since then over 50 Acts had made amendments to it with more than 400 individual amendment operations. A lawyer wanting to add that Act to his personal library would need to purchase from the Printing Authority of Tasmania the 1978 consolidation and each of the 50 Acts that had amended it since 1978. Then she would need to manually apply each of those 50 Acts in turn to the 1978 consolidation. But the PAT would not guarantee that all amendments were included in the package, and any amendments made consequential to another substantive Act were also unlikely to be included. This made it extremely difficult to verify the completeness of any particular consolidation.

In between published consolidations (or reprints), every repository of legislation (every law library including those of corporations, the University, and government departments, as well as public libraries) maintained their own copies of each substantive Act. Where having the most recent text of the law was important, changes were cut out of amending Acts and pasted in to the consolidations, and, in some cases (including OPC), were simply marked on the consolidation with pen. It was particularly important that the OPC maintain their own consolidations because Parliament needs to know the current text of a law before they can amend it. Because official reprints were falling behind, the state of these consolidations was rapidly deteriorating. One commentator notes that 'It is generally recognized that physical alteration of principal legislation to take account of textual amendments is tedious and, where there is extensive amendment, it may be in practice virtually impossible.'³ Figure 1 shows an example of a page of one of these unofficial consolidations. There are some legal publishers who still provide a service whereby clerks come into your library and apply all of the amendments of the previous parliamentary session using scissors and paste (but not in Tasmania any more).

² For more detail about different amendment strategies and generally about the EnAct system and associated research see Arnold-Moore (1998) *Information Systems for Legislation*, Ph. D. Thesis, RMIT.

³ Editorial (1990) "Textual amendment" 11 *Statute Law Review* iii-iv.

of 1998, 10% of households had Internet access.⁴ Approximately 20% of Australian adults accessed the Internet at some time during 1998⁵ and, since most schools and universities in Australia provide access to students, this figure is sure to be higher for younger Australians. Most public libraries in Australia also provide free access to the Internet. By making legislation available on the Web in addition to the traditional paper-based service, legislation is readily available to a whole new audience.

Web delivery of legislation also provides better access for those with disabilities. For those who have mobility problems, either through temporary illness or permanent disability, or the elderly, they can now access legislation in their own homes, rather than having to mail-order expensive and heavy paper copies or visit libraries with public access. Paper-based publications are also inappropriate for the sight impaired. By using Braille teletype machines or even voice synthesis, the electronic text of the legislation is more accessible.

The cost of this duplicated effort of maintaining legislation repositories is very difficult to measure but it is clearly substantial. The OPC seemed the logical place to concentrate the effort because this was the origin of the legislative material. Since the OPC also needed to maintain consolidations, it made sense to combine the responsibility for creating new legislation with the responsibility for maintaining consolidations that were of sufficient quality that all users of legislation could share them. Rather than simply providing the public with the legislation as enacted, ideally the new system would provide access to timely consolidations. This would mean the elimination of paste-ups altogether bringing great savings to all users of legislation who had to maintain their own consolidations.

Another important outcome of the proposed project was to improve the effectiveness and standing of the law and the Parliament. If it was hard for the legal community and the citizens to get an accurate view of the current law, it was also hard for the Parliament. Since the vast majority of legislation considered is amending legislation, it is vital to the Parliament to have access to up-to-date consolidations in order to see how proposed amendments change the existing law. The democratic process is also better served by broad community access to current legislation so that any affected groups can lobby their representatives appropriately.

All of these outcomes were sufficiently desirable to the Tasmanian people, that the Tasmanian government committed the considerable resources needed to a project to deliver all of these outcomes.

1.3. Planned out puts

The Legislation Systems Project began in 1992. Funding for the project was established with the following out puts to be delivered.

The first requirement was to develop and implement technology for drafting, managing and delivering legislation. This part of the project was put to tender and

⁴ Australian Bureau of Statistics.

⁵ Australian Bureau of Statistics.

RMIT together with Ferntree Computer Corporation⁶ were successful in the tender and produced the EnAct system. This included the public face of EnAct—the authorized public web site⁷—and also the drafting system behind the scenes. The more hidden side of EnAct was a system for the OPC to produce Bills and other documents for Parliament, and to populate and maintain the repository of legislation to which the public Web site gives access.

In order for this new system to be useful, the Tasmanian Statute Book needed to be consolidated in SGML form. This required a team of dedicated people sifting through the official and unofficial consolidations and keying them into SGML form, applying any intervening amendments as they went. An extensive, three stage proofreading regime was applied to these outputs so that the OPC could be confident that the electronic repository was correct.

Since the Tasmanian legislation production process was still based on paper technology with few of the drafters even using a computer, an extensive business process re-engineering process was required. The re-engineering effort was managed by the Corporate Information Projects Unit (now Whole of Government Projects Unit) to assist the technology-naïve OPC in dealing with the information technology providers and to provide strategies for managing the extensive change required as a result of implementing such a project. During re-engineering the new processes were captured in a workflow tracking system that forms part of EnAct.

Another outcome of this project was to provide legislation to establish the database as the authorized version of legislation. The *Legislation Publication Act 1997*⁸ provides for the electronic repository maintained by the OPC to be the official authorized version of the legislation. It gives considerable power to the Chief Parliamentary Counsel to maintain that repository and to publish authorized versions of legislation. This unprecedented step illustrates the confidence that the Tasmanian government had in the consolidation process and the EnAct system. At the time of writing Tasmania was still the only known jurisdiction to make the electronic repository normative.

2. What makes legislation different?

2.1. Complex regular structure

Legislation in Tasmania is similar in many ways to the legislation of the UK and most other jurisdictions that are former British colonies including other Australian jurisdictions, Canada and its provinces, and the United States. All Acts are divided into numbered sections (subordinate legislation variously calls these rules, regulations or clauses but the concept is the same). In larger Acts, these sections are grouped into

⁶ In 1996, Ferntree Computer Corporation <<http://www.ferntree.com.au/>> was bought by GE Capital Information Technology Solutions and the focus of their business moved away from document technology. EnAct and the underlying Structured Information Manager is now being marketed in Australia through Aspect Computing <http://www.aspect.com.au/Aspect/Products/products_sim.html> and <<http://www.simdb.com/>>.

⁷ <http://www.thelaw.tas.gov.au/>

⁸ [http://www.thelaw.tas.gov.au/summarize/s/1/?ACTTITLE=%22LEGISLATION%20PUBLICATION%20ACT%201996%20\(NO.%2017%20OF%201996\)%22](http://www.thelaw.tas.gov.au/summarize/s/1/?ACTTITLE=%22LEGISLATION%20PUBLICATION%20ACT%201996%20(NO.%2017%20OF%201996)%22)

numbered Chapters, Parts, Divisions, and Subdivisions. Each of these components has a heading, headnote or sidenote. Sections are further broken down into subsections, definitions, paragraphs, subparagraphs and other more specialized components.

In traditional paper publications, this structure is represented by typographic conventions. The headings of higher level components are in larger fonts and a mixture of bold and italic. Section headnotes are typically bold. Lower level elements are distinguished by a combination of multiple levels of indentation, and numbering conventions. Section numbers are typically in Arabic, subsection numbers in bracketed Arabic, paragraph delimiters in bracketed lower-case alphabetic, and subparagraph delimiters in bracketed lower-case Roman. Electronic delivery tools such as older subscriber dial-in services and even newer Web services have not been particularly faithful in reproducing this typographic information to the user, thereby obfuscating the structure present and devaluing the status of these electronic services. For widespread acceptance in the legal community, this structure must be supported.

Each of these sections is a logical unit that stands on its own but is related to other parts of the Act. There are extensive networks of cross-references between sections within an Act and also to components in other Acts. These can either be explicit text like "section 35(2) of this Act" or be more implicit like an occurrence of a defined term.⁹

Individual documents can contain as few as four sections or more than a thousand. This means that whole documents can be very large. The *Stamp Duties Act 1931(Tas)* is almost 1Mb of text, and the *Income Tax Assessment Act 1936 (Cth)* is almost 50Mb of text. For on-line viewing, these large documents must be broken down into manageable units without sacrificing the important context information of the surrounding text. Since all Acts contain sections, this provides a logical unit at which to fragment large documents. By displaying a table of contents along with the fragment, important context information remains available to a user. While this fragmentation approach contrasts the traditional delivery by the government printer of whole Acts, users of an electronic delivery system may also wish to print the whole Act rather than one fragment at a time.

The way around this contrast is to store legislation in a format that allows the logical structure of the document to be represented so that all of the typographic elements can be reproduced and fragments can be easily extracted for on-line delivery. SGML¹⁰ is an international standard that supports the representation of logical structure. It is widely used in the legal publishing field because of its suitability for legislation and complex legal commentary. Other advantages of SGML include the longevity of the

⁹ The AustLII collection captures these as hypertext links at the moment, but EnAct does not at present because AustLII automates this process and, although very helpful, such automation can misrepresent the scope of definitions where they are limited to particular components. Because EnAct is the authorized repository, such errors of scope cannot be permitted.

¹⁰ International Organization for Standardization. (1986) *Information processing – text and office systems – Standard Generalized Markup Language (SGML)* ISO/IEC 8879:1986. The newer Web standard, XML, which is a simplified subset of SGML, provides similar benefits although it was not available at the beginning of this project: World Wide Web Consortium. *Extensible Markup Language (XML) 1.0*, W3C Recommendation, 1998; Bosak and Bray (1999) "XML and the Second-Generation Web" 280 *Scientific American* 89-93.

data,¹¹ vendor independent tool sets,¹² the ability to validate the structure of each document against drafting standards,¹³ and the separation of the creation of content from the presentation of that content.¹⁴ SGML was the required encoding standard in the original Request for Tender for a legislation management system. By using SGML, Tasmania is able to automate the fragmentation of an Act, or the joining of a table of contents and set of fragments back into a single Act. This allows documents to be stored either whole or as fragments depending on retrieval requirements of the relevant legislative process stage.

2.2. Changing over time and the need to retain old versions

The most notable distinction of legislation, other than its complex structure, is that it varies with time. As discussed above, amendment legislation describes changes to the text of one or more Principal Acts. Either the amendments of the substantive law can be specified to commence at some time in the future or in the past. Amendments can apply before the substantive Act commences. They can expire reverting to the wording before the change was applied.

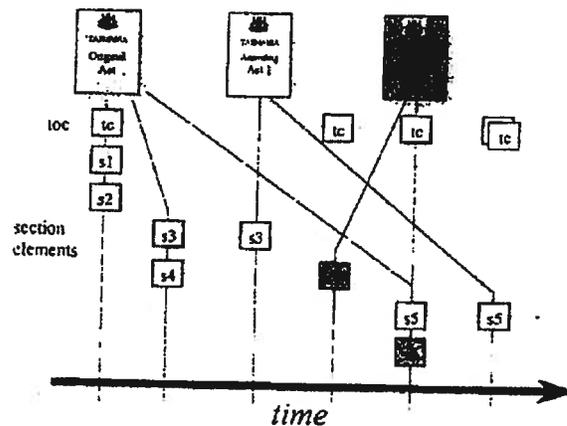


Figure 2: Example time-line of legislation.

¹¹ 'As an international standard, SGML ... is forced to be very stable. International standards are not changed whimsically. No single company or organization owns it and can steer it to their own advantage. Since SGML is so stable, software developers can safely build tools that employ the standard without the risk of having to constantly keep up with another company's dictums.' Travis and Waldt (1995) *The SGML Implementation Guide: A Blueprint for SGML Migration*.

¹² 'Users benefit from tools that work in predictable ways and can use off-the-shelf tools instead of having to develop their own validation and transformation programs.' *ibid*.

¹³ Arnold-Moore (1998) *Information Systems for Legislation* Ph.D. Thesis, Royal Melbourne Institute of Technology.

¹⁴ Barron (1989) "Why use SGML?" 2 *Electronic Publishing - Origination, Dissemination and Design* 3-24.

Figure 2 shows a time-line for a particular piece of legislation. An original Act commences sections 1 and 2 on an initial date, sections 3 and 4 on a later date, and section 5 on an even later date. After section 3 and 4 commence but before section 5, Amending Act 1 changes the text of section 3, and on a date after section 5 commences, substitutes a new section 5. The Amending Act 2 performs a retrospective amendment, repealing section 4 before Amending Act 2 actually commences. It also inserts a new section 6 at the same time section 5 commences. This time line shows each of the fragments and the table of contents and the lifetime for which each of these is valid. The first table of contents is not invalidated by a small textual change in section 3. A repeal, substitution, or insertion of a new section requires the table of contents to be updated.

2.3. User requirements – 'point-in-time' access to legislation

With these desired outcomes and these distinct properties of legislation in mind, it is important to design a user interface that supports these properties in ways that help users achieve their desired goals. Users require a variety of access methods to legislation.

While paste-ups and printed consolidations provide the current text of an Act, that does not solve the whole problem. The Renton Committee,¹⁵ when considering the needs for maintaining UK legislation as early as 1975, discussed the possibility of storing a complete historical record of the legislation as an alternative to the common method of simply maintaining the current version. Despite Canada's decision to use only the current file, they supported the historical file (back to the last complete consolidation) as the preferred option. Campbell and McGurk¹⁶ discuss the use of an information retrieval database of consolidations within a drafting office in 1987. These sources clearly indicate a need for keeping track of past versions of legislation.

There are two reasons for this. The first and primary reason is that legal researchers need to know what the law was at particular points in the past.¹⁷ A case goes before a court often many years after the relevant events took place but the court generally must apply the law at the time the disputed events took place, not the current law.

Even when the current law is being applied, previous versions of the law may still be relevant:

Exclusive reliance upon the reprint of regulations as amended ... sometimes, in a matter of interpretation, deprives the court of the advantage of seeing how the regulations were developed by amendment and why the amendments were made. ... It is not often that there is either need of or advantage in looking at the more authentic materials from which the Government Printer

¹⁵ *The Preparation of Legislation (The Renton Report)*. Cmnd 6053. 1975.

¹⁶ Campbell and McGurk (1987) "Revising statutes with computer support" 8 *Statute Law Review* 104.

¹⁷ Greenleaf *et al* (1995) "Public access to law via Internet: the Australian Legal Information Institute" 6 *Journal of Law and Information Science* and Robertson and Merrick. (1994) "Proposal for participation in the Workshop on Hypertext Systems and Version Support" in Durand *et al Proceedings of the Workshop on Versioning in Hypertext Systems*, pp. 35-38.

has reconstructed his convenient and perhaps more intelligible text. But this case happens, I think, to be such a one.¹⁸

While it is extremely unwieldy to provide access to every previous version of a particular piece of legislation in paper form, electronic delivery can solve the storage and presentation problems. What users want is the ability to specify a time point and then search and browse the collection as it was at that time point, viewing the law that applied at that time. While it is helpful to search the current time point and browse to previous versions, being able to search previous versions is particularly important where provisions have been repealed or replaced by provisions that cover different ground.

Since legislative documents are frequently very large, on-line retrieval needs to present fragments to the viewer.¹⁹ Care must be taken to preserve information about the context of the fragments and allow the user to navigate to related fragments with ease. In a hypertext environment like the Web, cross-references should all be activated as hypertext links.²⁰ Where these links are to legislation, they should be to the legislation as it was at the same time as the time point of the legislation being viewed currently.

It is also helpful, particularly for tracking changes in legislation to have history notes showing when the amendments were made and what Act effected them. Where these notes appear, it is desirable that they be linked to the actual amendment wording via hypertext.

Many legal practitioners (particularly OPC drafters) will be intimately familiar with the legislation collection or certain parts of it and require the ability to navigate straight to a known section of an Act.²¹ The search interface must allow search by Act number and year as well as title, and allow easy navigation to a known section. The browser interface must display a list of Act titles for the user to select the appropriate Act.

While on-line browsing demands access to individual fragments, a user that wants to print the whole Act does not want to navigate to each fragment in turn and print them separately. Even if the Web repository is not an authorized version, it is desirable to allow the user to download a whole Act so that it can be printed as a single unit.

¹⁸ *O'Neill v. O'Connell* (1946) 72 CLR 101, 122 per Dixon J
<http://www.austlii.edu.au/au/cases/cth/high_ct/72clr101.html>

¹⁹ This can also improve retrieval performance, see Arnold-Moore (1998) *Information Systems for Legislation*, Ph.D. Thesis, RMIT.

²⁰ Corbett (1992) "Indexing and searching statutory text" 84 *Law Library Journal* 759-67; Hoey (1988) "The discourse properties of the criminal statute" in *Wagner (ed) Computer Power and Legal Language*; and Schweighofer and Scheithauer (1996) "The automatic generation of hypertext links in legal documents" in Wagner and Thoma (eds) *Database and Expert Systems (DEXA '96)*.

²¹ Tapper (1970) "Computers and legislation" 23 *Alabama Law Review* 1-42.

3. How does EnAct fulfill these requirements?

3.1. The public web site

The EnAct repository makes use of the Structured Information Manager to store SGML fragments and associated metadata directly. Acts are stored in SIM as fragments with a timestamp marking the time interval (start and end time) over which the fragment or table of contents is valid. The SIM repository incorporates an SGML parser allowing sophisticated indexing based on the logical structure of the SGML fragments. Time point searching is also supported. When a time point is specified, a filter is applied to the database leaving only those fragments and tables of content that were valid at the specified time. This allows a snapshot of the database to be searched and browsed as if it were the entire database.

The SIM web server uses the same SGML parser. This allows HTML to be generated dynamically from the fragment repository for delivering Web pages. This means that the same repository can be viewed in a number of different ways. In the current public site one can view a fragment with or without history notes. A single fragment can be viewed on one side with a table of contents on the other, or with the table of contents and all of the fragments united together into a single document.

All cross-references are activated as hypertext links. Because all hypertext links are activated using queries in the database, it is just as easy to do reverse hypertext links—i.e. show a list of all fragments that refer to this fragment or this Act. When history notes are displayed, those that refer to amending Acts on the system are also hypertext links to those Acts. The table of contents can be viewed either as a conventional section number and headnote or with a section number and list of links to successive versions of the corresponding fragments. This provides a quick overview of the history of a provision. A new feature is a previous and next version button for each fragment.

3.2. On-demand printing

In addition to the public web site, which delivers HTML to user's browsers, the Printing Authority of Tasmania also uses the same repository to generate 'on-demand' authorized reprints. A member of the public can contact the PAT shop and request a reprint of a particular Act on a specified day. They search the repository to find the relevant table of contents and fragments at that time point, which are joined together into a single SGML document that is translated into Rich Text Format using SIM formatters. This RTF representation is then passed to high-speed printers for paper output. Each authorized version is preceded by a certificate of authenticity from the Chief Parliamentary Counsel.

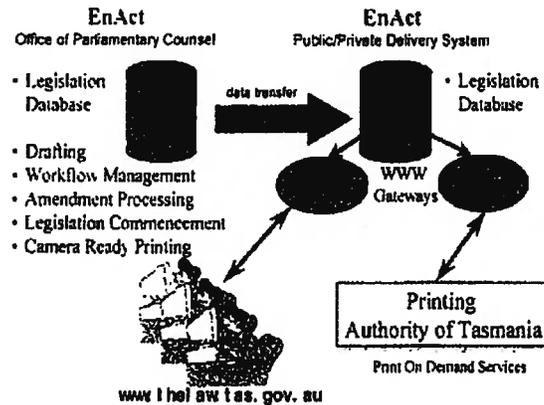


Figure 3: The EnAct repositories and how they are accessed

At the end of the document two tables appear. The first shows the name and the time of commencement of all amending Acts applied to that Act since it was enacted. The second lists each provision that has been amended and how it was amended.

The database that is shared by the public Web gateway and the private Printing Authority gateway is not the only repository in the EnAct system. A working database is also kept in the OPC. This database contains everything in the public repository, but also contains the politically sensitive draft Bills in preparation, and all of the workflow information about the status of Bills and other draft legislation. Data is periodically migrated from the production database to the publication database providing a level of protection from intrusion and preserving the integrity of the data on the public repository.

3.3. Behind the scenes – the drafting environment

In the past, any consolidation has been a laborious manual process. Drafters or clerical staff were required to sit down with the text of an Amending Act and go through it, section by section, applying each amendment to one or more Principal Acts. In the past drafters composed amending legislation directly, combining the process of deriving the effect of the amendments with the text describing them. The Tasmanian government presented RMIT with a different approach. The drafters would mark amendments directly on a consolidation of the Principal using strike-through and underline markings familiar to many lawyers. Amendment wordings for those markings would then be generated automatically. For example, to draft an amendment to section 74F of the *Racing and Gaming Act 1952*, a drafter would check out the appropriate version of the Act from the repository as in Figure 4. Then each change would be marked on this version using strike through and underline giving a version like that in Figure 5.

Principal Act - Racing and Gaming Act 1982

Obligations of licensees
 74F. A licensee -

- (a) must not make bets except in accordance with the licence; and
- (b) must not, as a licensee, 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
- (c) 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) 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
- (e) 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 4: Example section.

These markings are then captured in an internal (SGML) representation of the changes called a Change Description Document (CDD). These changes are then used to generate amendment wordings, which are appended to a stub or substantive Bill as in Figure 6.²²

This process is managed by a workflow enactment service that keeps the CDD and generated amendments together so that, if and when the amendment Act commences, the amendments can then be automatically applied to the principle to generate new fragments in the historical repository.²³

Markup of Principal Act - Racing and Gaming Act 1982

Obligations of licensees
 74F. A licensee -

- (a) must not make bets ~~except in accordance with the licence;~~ *Substitute Text* and
- (b) must not, as a licensee, ~~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~~ *Substitute Element*
- (c) ~~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~~ *Substitute Element*
- (d) ~~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~~ *Substitute Element*
- (e) ~~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.~~ *Substitute Element*

²² This process is described in greater detail in Arnold-Moore (1997) "Automatic generation of amendment legislation" in *Proceedings of the International Conference of Artificial Intelligence and Law (ICAIL'97)*.

²³ More detail on the workflow enactment aspect of the system can be found in Arnold-Moore (1993) *Information Systems for Legislation* Ph.D. Thesis, RMIT.

Figure 5: The same fragment with amendments applied.

4. Where to from here?

Because we see structured documents all the time, we are used to inferring structure from typographic cues in the printed page. Largely these cues can be reproduced in electronic forms.²⁴ However, there is little paper-based precedent for representing multiple versions of a single document or part of a document. RMIT in collaboration with the Tasmanian government are working on ways of visualizing the temporal information contained within the EnAct repository, seeking ways of improving the interface and making the most of the information that has already been captured.

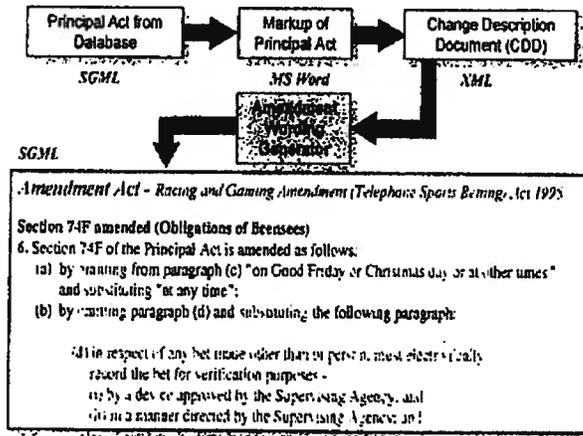


Figure 6: Using the CDD to generate amendment wording.

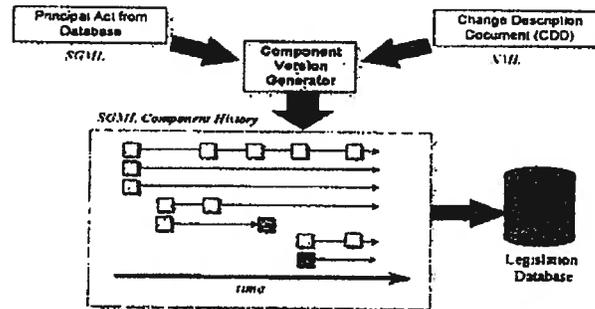


Figure 7: Using the CDD to generate fragments in the repository.

In addition to the automatic generation of amendment wording, techniques have been devised to automate the understanding of amendment text, generating CDD's from

²⁴ Bosak and Bray (1999) "XML and the Second-Generation Web" 280 *Scientific American* 89-93.

amending Acts.²⁵ The incorporation of these techniques into EnAct would enable the simplification of business processes to support the amendment of amending Acts, and the production of Parliamentary amendments (amendments proposed for debate in Parliament).

Work is currently being undertaken to develop a Web-based workflow enactment service and document management system using the Structured Information Manager as the basis of the repository. This would enable the EnAct workflow system to be moved away from a custom "thick" client to a "thin" Web client. This facilitates the movement of workflow tasks outside of the OPC, to Parliamentary Clerks and Cabinet Secretaries, who have no drafting responsibilities but can record changes in status of Bills as they happen.

The lessons learned in the Tasmanian context and the solutions developed here are applicable to other jurisdictions. We are constantly working to implement this technology more widely and assist other jurisdictions to satisfy their public duty obligations to provide the highest quality of access to legislation that is reasonably possible. If a jurisdiction as small as Tasmania can provide such a service, how much more so can larger jurisdictions with considerably greater resources fund similar projects.

The Tasmanian government is leading the world in taking up its responsibilities to provide public access to legislation. We have put considerable effort into understanding the distinct properties of legislation and meeting the needs of the users of legislation. By centralizing the responsibility for maintaining consolidations, we have eliminated re-keying and saved business and government considerable costs in maintaining their own consolidations. By locating that responsibility in the same office that drafts new legislation, we have enabled unprecedented automation of the drafting of legislation and the its consolidation. Because the consolidation is automated, any delay in applying amendments to up-date the repository can be eliminated. Innovative use of the Internet has delivered a service far beyond what was possible with paper-only publication.

Bibliography

- Arnold-Moore, T. "Automatically processing amendments to legislation" (1995) in *Proceedings of the International Conference of Artificial Intelligence and Law (ICAIL'95)*.
- Arnold-Moore, T. "Automatic generation of amendment legislation" (1997) in *Proceedings of the International Conference of Artificial Intelligence and Law (ICAIL'97)*.
- Arnold-Moore, T. *Information Systems for Legislation*, Ph.D. Thesis, RMIT, 1998.

²⁵ Arnold-Moore (1995) "Automatically processing amendments to legislation" in *Proceedings of the International Conference of Artificial Intelligence and Law (ICAIL'95)*.

- Australian Bureau of Statistics Press Release, 29 May 1999.
- Barron, D. "Why use SGML?" (1989) 2 *Electronic Publishing – Organization, Dissemination and Design* 3-24.
- Bosak, J. and Bray, T. "XML and the Second-Generation Web" (1999) 280 *Scientific American* 89-93.
<<http://www.sciam.com/1999/0599issue/0599bosak.html>>
- Campbell, C. and McGurk, J. "Revising statutes with computer support" (1987) 8 *Statute Law Review* 104.
- Corbett, M. "Indexing and searching statutory text" (1992) 84 *Law Library Journal* 759-67.
- Greenleaf, G. *et al.* "Public access to law via Internet: the Australian Legal Information Institute" (1995) 6 *Journal of Law and Information Science*
<<http://www.austlii.edu.au/austlii/libs/paper.html>>
- Hoey, M. "The discourse properties of the criminal statute" (1988) in Walter (ed) *Computer Power and Legal Language*
- International Organization for Standardization *Information processing – text and office systems – Standard Generalized Markup Language (SGML) ISO/IEC 9979:1986.*
- Robertson, J. and Merrick, F. "Proposal for participation in the Workshop on Hypertext Systems and Version Support" (1994) in Durand *et al Proceedings of the Workshop on Versioning in Hypertext Systems*, 35-38. <<ftp://bush.cs.tamu.edu/echt/vers-wkshp/VWReport.A4.ps.gz>>
- Schweighofer, E. and Scheithauer, D. "The automatic generation of hypertext links in legal documents" (1996) in Wagner and Thoma (eds) *Database and Expert Systems (DEXA '96)*.
- Tapper, C. "Computers and Legislation" (1970) 23 *Alabama Law Review* 1-42.

Travis, B. and Waldt, D.

The SGML Implementation Guide,
1995.

World Wide Web Consortium

Extensible Markup Language (XML)
1.0, W3C Recommendation, 1998.
<<http://www.w3.org/TR/REC-xml>>

Editorial

"Textual amendment" (1990) 11
Statue Law Review iii-iv.

U.K. Command Paper

The Preparation of Legislation (The
Renton Report). Cmnd 6053, 1975.