THE GILBANE REPORT
on Open Information & Document Systems

STATE OF WISCONSIN LEGISLATURE TEXT 2000 — REENGINEERING FOR DOCUMENT MANAGEMENT

We begin our first issue of volume 2 with an examination of a sophisticated document management implementation — the TEXT 2000 project designed and implemented by the State of Wisconsin and Andersen Consulting.

What makes this project especially instructive is that it illustrates many of the key concepts we associate with compound document management solutions: reengineering the document process to create more efficient workflows, building an information/document repository that allows for document information to be shared instead of being re-created, and providing for both electronic and printed delivery of documents in multiple formats.

The TEXT 2000 project is also interesting because most of the complex document management implementations we hear about involve technical (engineering or product) data. This example clearly shows how complex a process it can be to manage "non-technical" documents that need to be managed by administrative, legal and financial parts of a corporate or government enterprise.

We think you will find the integration of a variety of different document management software applications described here fascinating and thought provoking. The goal was to match the best off-the-shelf solutions available to the specific requirements of TEXT 2000, regardless of which vendor supplied it.

We are grateful to Andersen Consulting, especially the project team in Wisconsin, for spending the time and effort to provide us with a write-up of the TEXT 2000 project. We also thank the State of Wisconsin for allowing us to publish the case study.

CONTENTS

STATE OF WISCONSIN LEGISLATURE TEXT 2000 —

REENGINEERING FOR DOCUMENT MANAGEMENT

EXECUTIVE SUMMARY

The Existing Legislative Process

• The State of Wisconsin and its five supporting agencies are responsible for drafting, maintaining, and distributing all legislative documents. The process of drafting reviewing, updating and delivering the various types of documents is complex and difficult to manage effectively.

• Different legislative staff and attorneys need to continually share documents throughout the process of creating and updating documents.

• Maintaining the quality of the documents while keeping up with the increasing quantity had become concerns.

The Document Management Solution

• A document reengineering process pointed to a number of ways to improve workflow and to eliminate tasks that did not add value.

• The most significant non-value-added tasks were related to the recreation and re-invention of existing information. It was determined that a central document repository with appropriate controls and workflows could eliminate these tasks.

• Implementing a complete solution that used the best available technology required integrating relational database, document management, document publishing, and electronic viewing products. from different companies in a client/server environment.

• The business case involves both immediate benefits such as improved throughput, quality, usability and flexibility, and long term benefits including future compatibility, scalability, extensibility, and support for future public access.

Lessons Learned

• The requirement to handle structured documents needs to be balanced with the demands for hiding tagging complexity from certain types of users.

• Defining the document objects and object classes requires careful analysis. Too few classes can limit flexibility, and too many classes can cause the system to be difficult to use and maintain. It pays to establish clear guidelines for determining how many object classes are necessary.

• Technical integration issues are complicated by needs to integrate multiple application solutions across multiple platforms. Products with well-defined, extensible, and open APIs are critical for success.

• Users will be less threatened of new document management systems if they are included in the planning process early on, provided with non-threatening training environments, and given some responsibility for ongoing support.

• The organizational impact of process reengineering and implementation of a new workflow with new technology can be significant. This needs to anticipated and managed from very early on in the project.

INTRODUCTION

Because they model how we do business—and because no two businesses are identical—document management systems vary in the way they are implemented. Still, there is much to be learned from studying how such systems are applied in different circumstances. Businesses that are considering implementing a document management system would do well, for example, to learn from the experiences of those government agencies that have applied such systems against their often-critical document management needs.

The Wisconsin Legislature provides an interesting case study in "mission-critical" document management. Its new legislative Document and Workflow Management application, TEXT 2000, was developed with Andersen Consulting to address the same kind of document quality and productivity issues faced by organizations today. In this report, Andersen reports on the major functional and technical components of this industrial-strength document management system, and describes the lessons it holds for systems of similar size and complexity.

The State of Wisconsin Legislature and its supporting agencies are responsible for drafting, updating, and distributing all documents generated by the legislature. The document processes begin with the legislative drafting process and end with updated State statutes. Attorneys and legislative support staff prepare and revise legislative drafts and other legislative information.

The challenge facing the Legislature was how to maintain the quality of a steadily increasing quantity of bills without impeding their progress through the legislative process. Until recently, the system used to support their document processes was an outdated mainframe application. This application and the document processes it supported were not producing drafts quickly enough; quality was also steadily decreasing under the growing volume of bills filed. Support for the State's existing application was about to end when the State began to seek a new document management solution.

TEXT 2000, the solution that the State decided to implement, employs a unique client/server architecture based upon commercially available off-the-shelf software (COTS) components and a graphical user interface (GUI). Proposed by Andersen Consulting, the system's two key software components are Interleaf 5 and Documentum: Interleaf to support the drafting and publishing of documents, and Documentum to store and maintain all the key documents used by the legislature and its agencies. Sun Microsystems provided the Database and application server components, while Intel-based computers running Microsoft Windows are used as client workstations. The Andersen project team integrates these products and develops custom functionality to give the State leading-edge document management capabilities.

A pilot application of TEXT 2000 is now in production at a legislative agency, the Revisor of Statutes Bureau, which now produces two key documents using TEXT 2000. Benefits recognized in the Revisor's Bureau to date include increased quality, faster turnaround and reduction of external typesetting costs. Custom validation routines now flag errors that were difficult to catch before. The elapsed time to produce one of these documents is expected to drop from about four weeks to two-and-a-half weeks. Additionally, the Revisor now produces camera ready copy in-house, saving the State from external typesetting charges which average $35 per page.[1]

---

[1] During 1992, 5,865 pages were produced externally at a cost of approximately $205,000.

THE EXISTING LEGISLATIVE DOCUMENT PROCESS

In Wisconsin, five major agencies support legislative activities: the Legislative Reference Bureau, Revisor of Statutes, Senate Chief Clerk, Assembly Chief Clerk, and the Legislative Council. The Senate and Assembly Chief Clerks manage the legislative process within chambers; the Legislative Reference Bureau drafts and maintains bills; the Legislative Council provides drafting and research activities; and the Revisor of Statutes updates and publishes State Statutes and administrative code.

The legislative document process begins with the Legislative Reference Bureau (LRB), the agency responsible for drafting legislation. Requests for legislative drafts (bills, resolutions and amendments) are assigned to a drafting attorney who works with the requester to determine exactly what the proposed legislation is meant to accomplish. The attorney must then research the statutes to determine which portions must be changed to accurately implement the proposal as legislation. The attorney must also check for statutory cross-references and other legislation that may conflict with the draft. Much of this research today is manual, and attorneys tend to write drafts by hand because the present system is so hard to use. The only way to find current legislation by subject or author is to search an index-card file maintained by the LRB. Although statutes and some larger bills can be located electronically using so-called *full-text search* capability (inverted index search), most other documents such as memos and the Administrative Code cannot.

> *"Much of this research today is manual, and attorneys tend to write drafts by hand because the present system is so hard to use."*

After an attorney has completed the preliminary draft, an editor reviews it for grammar, style, clarity, proper use of terms, etc. The editor also re-checks statutory cross-references and tries to identify any other legislation that may conflict with the draft. If no significant changes are required, the draft then goes to a word processing operator (WPO) for entry into the system. If significant changes are needed, the editor gives the draft and comments back to the attorney, who in turn makes all necessary changes and then submits the revised draft to the WPO. This process continues until the attorney is satisfied with the draft. During busy legislative periods, many different drafts are circulated, and legislative in-baskets tend to pile up with work. At any point in this process, the only person who really knows the status of a proposed bill is the one who has the most current draft.

Once a draft is approved, it is introduced as a proposal by the Chief Clerk of the house of origin of the bill. It is then referred to committee by the house leader for review. The committee generates a report on the proposal, and the house leader then schedules the proposal for debate on the house floor. The assembly enters information in a computer on the floor, and both houses also take notes manually. This information is used to generate the *Daily Journal* and *Bulletin of Proceedings*. These documents are manually compared to ensure the accuracy and consistency of the information. Much of this information is rekeyed later to publish a weekly summary called the Bulletin of Proceedings.

If the proposal passes both houses without amendments, it goes back to the Legislative Reference Bureau for *enrollment.* Resolutions need no further action; approval by the legislature (enrollment) is all that is required. Bills, however, must be approved by the Governor. If the Governor enacts (signs) the bill, it becomes an Act. The Legislative Reference Bureau then publishes the Act and gives a copy to the Revisor of Statutes. The Revisor of Statutes is responsible for updating the Wisconsin Statutes to include the changes specified by the Act.

THE SOLUTION

Document Reengineering

One of the unique aspects of the Andersen approach was its emphasis on document process reengineering. The goal of a process-reengineering effort was to derive more benefits in throughput, flexibility, quality and usability than would be realized by simply introducing new technology. To achieve these results, the project team analyzed the legislative document processes and streamlined some processes and totally reinvented others. During the reengineering effort, the project team looked for ways to:

- Eliminate non-essential tasks
- Consolidate or re-arrange tasks to make them more efficient
- Automate manual and time-consuming tasks
- Better manage document processes
- Enable new processes

The biggest benefits were realized by eliminating so-called "non-value-added" tasks. By far the biggest non-value-added task was the re-invention and recreation of information. Documents originating in other agencies had to be retyped for lack of a way to transfer the information electronically. Searching methods were rudimentary, and research would often be repeated on a subject because a particular document on that subject could not be found. Also, many documents within the State shared information, and information originating in one document would be re-keyed in several others. The new processes eliminate rekeying information.

Large amounts of time were also spent on activities such as filing and distributing paper documents, printing (or "playing out") documents so that users could see what they actually looked like, and maintaining files of index cards with document-related information. With new technology, many of these tasks could be eliminated.

As document processes were reengineered, it became apparent that different processes would need to change in different ways. Some processes became more streamlined, while others were totally re-invented. The fundamental legislative drafting process stayed the same, but by eliminating non-value-added tasks and making other process improvements, the overall production cycle time was greatly reduced, throughput and turnaround time increased, and quality improved.

In contrast, the Senate and Assembly Chief Clerk offices' document processes were totally re-invented. The documents produced by these agencies share information, but their processes centered around the documents instead of around points of information. As mentioned above, data was captured separately for both the Daily Journal and the Bulletin of Proceedings, although much of the information was the same. The Weekly Schedule and the Daily Calendar also use much of this information, but once again these documents were each created separately.

The Chief Clerks' document processes were radically redesigned around the information used to create the documents. Under the new processes, floor actions are recorded directly and stored in the database. Information is also extracted ("shred") from other documents and is stored in the database. The resulting documents are then produced from the database automatically, using database publishing techniques.

Before and after views of the Chief Clerks' document processes are depicted in the following workflow. By capturing the information and re-using it to produce certain documents automatically, the Legislature significantly reduced the total number of production steps required to produce and maintain them.

"By far the biggest non-value-added task was the re-invention and recreation of information."
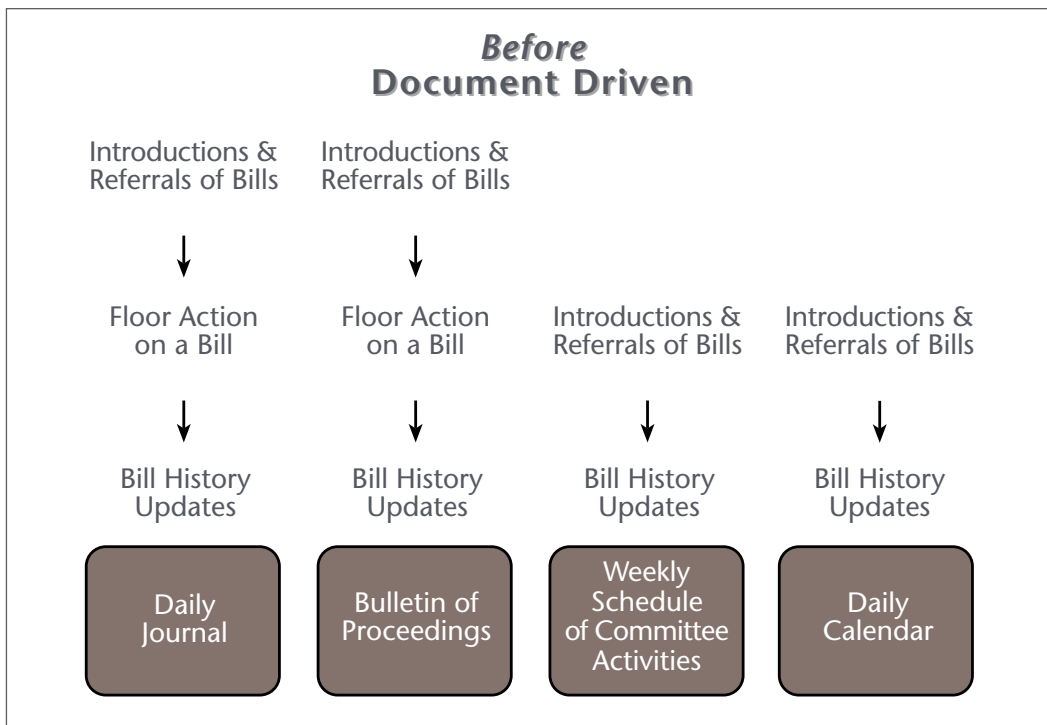
Figure 1

The Original

Document

Development

Process

Before
Document Driven

Introductions & Referrals of Bills	Introductions & Referrals of Bills		
↓	↓		
Floor Action on a Bill	Floor Action on a Bill	Introductions & Referrals of Bills	Introductions & Referrals of Bills
↓	↓	↓	↓
Bill History Updates	Bill History Updates	Bill History Updates	Bill History Updates
Daily Journal	Bulletin of Proceedings	Weekly Schedule of Committee Activities	Daily Calendar

The Document Management Solution

The State's legislative document processes required a system with the following document management functions:

Repository: Storage and access of documents from multiple distributed locations.
Drafting: Entry of text into the system.
Retrieval: Search and retrieval documents based on metadata and content.
Workflow: Management of the document production process.
Composition: Creation of documents from existing information.
Conversion: Conversion of existing documents to formats used in the new system.

Repository

A key requirement for TEXT 2000's document repository was that it be capable of handling very large volumes of documents. The Legislative Reference Bureau alone produces approximately 20,000 drafts each session. Besides the multiple versions required of each draft, in some cases several other reference documents must be stored with the draft. To ensure that enough document-history information is on hand for support legislative research, several sessions worth of drafts must be stored in the repository; Copies of old legislative drafts are also used as templates for use in drafting new bills.

Due to the location of the agencies and the users of this system, distributed access to the repository was also required. This requirement led to adoption of a client-server–based repository architecture. Besides meeting the access requirements, the client-server approach allows the system to be expanded easily to serve more users and agencies (and, potentially, for public access sometime in the future).

Since no one product meets all the State's needs, it was important that all tools selected, especially the repository, be extensible and easy to integrate with other tools. The repository had to (1) have an open applications programming interface (API), (2) be extensible in terms of data structure and behavior modification, and (3) be independent of other tools and file formats.

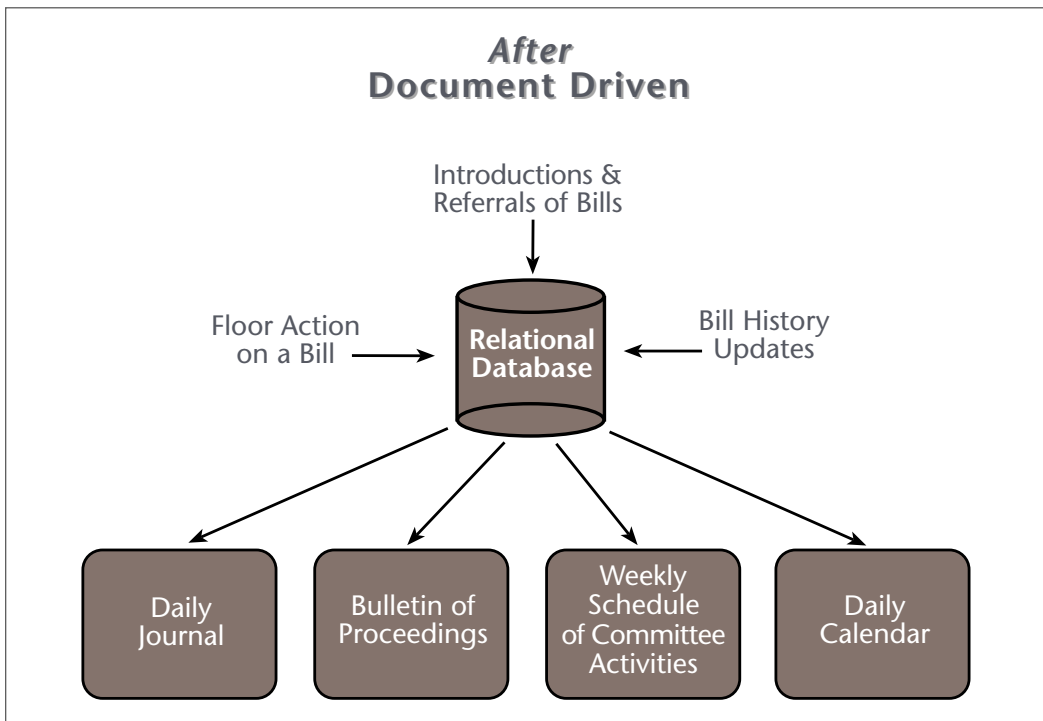*"The Legislative Reference Bureau alone produces approximately 20,000 drafts each session."*

After
Document Driven

Introductions &
Referrals of Bills

Floor Action
on a Bill → **Relational Database** ← Bill History
Updates

| Daily Journal | Bulletin of Proceedings | Weekly Schedule of Committee Activities | Daily Calendar |

The State also had special requirements for compound documents. In some cases all it required was for related documents to be grouped or associated together, such as documents related to a particular legislative draft. Other documents, such as the statutes, are, in fact, published together and thus require more complex "book" technology.

True book technology is more than simply organizing a group of documents in a hierarchy. Users must be able to manipulate a single document in a book, or work with the entire book as a whole (*e.g.*, printing). The individual documents within a book must be sensitive to their location in the book, as well as the other documents in the book. For example, changing the order of documents within the book should also change page numbers, chapter numbers, references, table of contents, etc., for all documents in the book.

Due to these requirements (and others addressed below), the State decided to use a Documentum repository. The current configuration for TEXT 2000 includes 32 Gigabytes of disk storage space. Documentum will eventually support a CD-ROM jukebox interface.

Documentum is a client-server–based, object-oriented document management product that offers more than just document repository capabilities. One of the most important features from an integration standpoint is its extensibility. Once object classes[2] have been defined for different types of documents in the system, the *behavior* of these object classes can be modified. A behavior is what happens to an object when a user performs an action on the object. For example, telling the system to print a document is an action, and sending the document to the printer is the behavior. In TEXT 2000 this behavior has been modified for Legislative Drafts (which have their own object class) so that the document is first sent to a program that places line numbers in the document before sending it to the printer.

Beyond extending object classes and behaviors, Documentum provides an open API to allow integration with other tools. The Documentum server has an API library for both C/C++ and the *gawk* scripting language. A C/C++ DLL is also available for Windows client applications, and the Documentum Workspace provides a full API through DDE.

> *"The State also had special requirements for compound documents."*

---

[2]A an object class is a type of object as opposed to a specific object. Legislative drafts are an object class, and a particular legislative draft is an object.

Drafting

One of the goals of TEXT 2000 was to make drafting (text entry) easier, so more information could be entered at the point of origin. The previous system was cryptic and hard to use, so much so that many users, such as drafting attorneys, hand-wrote or dictated legislation rather than using the system to input it themselves.

The drafting tool also had to possess powerful document publishing capabilities. Some documents produced by the State have complex page layout and formatting requirements; others call for rotated text or graphics. The ability to create complex graphics, charts and text layouts from a single tool was a high priority.

Requirements for ease of use and for document-publishing capabilities had to be balanced against a need to effectively tag information in documents for use by the system. Several documents produced by the State serve as sources of data that is extracted for use in other documents. As with other tools selected for the project, the drafting tool had to be extensible and it required an open API set to allow integration with other tools.

Interleaf 5 was selected as the TEXT 2000 drafting tool. A powerful WYSIWYG document publishing system which allows text to be tagged through the use of component markup, Interleaf also possesses the document publishing and compound document capabilities the State required and it provides an easy to understand user interface. With Interleaf 5, templates can be created for documents with complex formatting and tagging requirements, thus allowing even less experienced users to create complex documents.

Interleaf 5 is fully extensible. In addition to simple macros, it has a complete API and custom LISP programs can be created to access these API calls. With this technology, several custom routines were created by Andersen Consulting to extend Interleaf's capabilities and modify standard behavior to meet the State's requirements.

Interleaf also features a regular-expression–based file-processing tool, Cloverleaf, which was used extensively to create document filters. Some filters were used to import different documents into Interleaf format. Others were used to re-purpose or reformat a document, such as changing a Bill into an Act. Still others were used to create entirely new documents from existing documents, such as creating a list of all statutory cross-references using the statutes as input.

In addition to Interleaf, two other add-on tools were included in the system: Oracle Coauthor, which provides additional proofing and editing capabilities; and Smartleaf Compare, which is used to compare two different documents and highlight any differences.

Search and Retrieval

Because so many documents will be stored in the repository, a robust method is required for locating and viewing documents there. To meet the needs of the State, the search engine had to be able to locate documents based on ideas and on the attributes associated with these ideas. Once these documents were located, they had to be presented to the user in a read-only, WYSIWYG format.

As mentioned before, Documentum stores all documents as objects and associates with each document a custom list of attributes (based on object/document class). Documentum also provides a query language, DQL (Document Query Language), which is a superset of SQL.

In addition to its metadata searching capabilities, Documentum comes bundled with the Verity full-text search engine. Full-text searching, also supported by DQL, allows documents to be located based on both ideas and attributes. For example, a user could pose a

"One of the goals of TEXT 2000 was to make drafting (text entry) easier, so more information could be entered at the point of origin."

query to Documentum to locate all bill drafts that were requested by a democratic senator and which have the word "crime" and "reduction" in the same sentence.

Documentum then returns a list of all documents that meet the search criteria. The listed documents can then be viewed by the user, but only in a read-only, WYSIWYG format. The viewing tool, Adobe Acrobat, consists of three pieces; Distiller, Exchange, and Reader. The Distiller is used to transform a PostScript file into a Portable Document Format (PDF) file. The PDF file is used by the Reader to view and navigate the document. Acrobat Exchange is capable of creating hyperlinks and collating documents. All three are being used in TEXT 2000.

Using the extensibility of Documentum, the transformation of documents from Interleaf to Adobe format is invisible to the user. When a document is returned from a query and selected for viewing, Documentum uses the selected document to create a PostScript file, and from the PostScript file a PDF file. These files are stored with the original document. The PDF file is returned to the workstation, and the Acrobat Reader is launched with the returned file. Subsequent requests to view this document will access the PDF file stored with the document, which significantly improves the response time.

Workflow

The State's document production processes typically involve many steps and several different users. Under the previous system, all documents were printed and physically routed to each successive worker in the production process. The ability to eliminate manual steps and capture workflow information was important to the State.

TEXT 2000's workflow component is based upon a Documentum feature that allows the definition of standard workflows, called *routers.* For example, the standard workflow for a legislative draft consists of such steps as *drafting, editing, typing,* and *submitting.* The router for such this type of document defines a separate "task" for each of these steps.

The workflow component allows users to determine easily what work needs to be accomplished in any given day, and it eliminates manual distribution of paper copies. For each task, information such as begin time, end time, user performing the task, etc., is also captured. Capturing such data allows the system to create reports on the production process so that the State can identify bottlenecks and distribution of workload.

Composition

Composition involves creating documents from existing information. Interleaf and its Cloverleaf programming language provide ways of doing this, but the State required more capability here.

Often key pieces of information are created in one legislative document, but need to be replicated in others. An example of this is the *relating clause* in a bill. Although created in a bill, the relating clause also appears in the Daily Journal, the Bulletin of Proceedings, and the Special Order of Business document.

Wisconsin's solution to this problem was to "shred" data to the Documentum database and to republish data by creating other documents automatically. Smartleaf, a plug and play extension of Interleaf that is fully compatible with Documentum's Oracle database engine, was chosen to fill this role.

With Smartleaf, the original document-specific text, such as the relating clause, is tagged using ordinary Interleaf markup. When the user checks a document into the repository after editing, the relating clause is "shred" out of the draft and stored in the database by a unique identifier. When the time comes to produce another document, such as the

Daily Journal, Smartleaf retrieves this data and creates a new document (or adds to an existing one) using the data from this database.

Database publishing eliminates the need to re-key information and helps reduce the number of errors introduced by that process. It also dramatically decreases the amount of time required to create a document based on shared information. In addition, it provides access to bill status information in a relational database for producing reports on bill status, committee activity, and other related topics.

Conversion

The State plans to have TEXT 2000 in full production at the beginning of the September 1994 legislative session. In terms of conversion requirements, State documents fall into one of four categories: "as needed," "session-oriented," "continuous," and "historical."

"As needed" documents are created on demand, as required. Examples include legislative drafts, memos and letters. These documents generally do not require conversion.

Session-oriented documents are created only at the beginning of a legislative session. Information is added to these documents until the end of the session, at which point a new document is started. These documents generally do not require conversion. An example is the Bulletin of Proceedings.

Continuous documents are on-going or "living" documents. They are only changed— never created new. These documents require conversion to the new system. An example is the Statutes.

Historical documents were created originally as either as-needed or session-oriented documents. Once completed, however, they must be maintained for historical and research purposes. These documents require conversion to TEXT 2000. Examples include important legislative drafts from previous sessions.

Documents under the current system were created with ATMS, a mainframe word-processor. Data Conversion Laboratories (DCL) was chosen to perform the conversion of documents from ATMS to Interleaf because of their experience with this type of conversion.
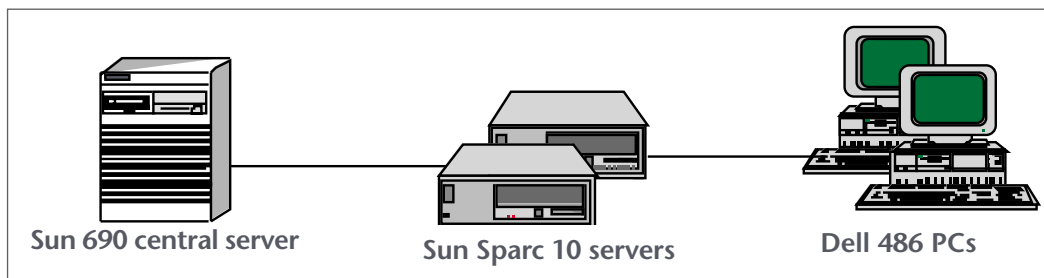

Figure 3.

The

Client/Server

Architecture.

Sun 690 central server   **Sun Sparc 10 servers**   **Dell 486 PCs**

Technical Architecture

TEXT 2000 is built on an open systems client/server architecture. A diagram of the architecture is shown below:

The hub of the system is a central data server that is used to store all documents and run all background processes. The central server is a Sun SPARCServer 690 with 256 megabytes RAM and 32 gigabytes of disk storage. The server is upgradeable to 1 gigabyte of RAM and 52 gigabytes of disk storage. Documentum, Oracle, Acrobat Distiller and custom batch routines created in C/C++, *gawk* and Cloverleaf all reside on the data server.

Attached to the data server are individual Token Ring hubs consisting of an application server and, in most cases, five or six workstations. The application servers run Interleaf and any other software that assists in the drafting process. Isolating groups of users on separate Token Ring hubs minimizes network traffic and ensures that each group always has enough resources to run its applications.

The application servers are Sun SPARC 10 model 51 workstations with 128 megabytes of RAM and 2.1 gigabytes of disk space. The Interleaf, Smartleaf Database Publishing, and custom interactive routines created in Cloverleaf, C/C++, and Lisp reside on the application server.

The workstations are Dell 486/66 PCs with 8 megabytes of RAM and 170 megabytes of disk space running Windows 3.1. The workstation serves as the interface to all other components of the system. The Documentum Workspace is used to find, access, and manipulate documents in the repository. Interleaf, used to edit documents, runs in an X-Windows session on the application server. Acrobat Reader also resides on the workstation, along with custom routines written in Visual Basic and C++.

BUSINESS CASE

The business case for TEXT 2000 can be broken into immediate benefits and longer-term benefits. Immediate benefits are those that improve the State's document-management processes right away, in such areas as throughput and turnaround, quality, usability, and flexibility

Longer-term benefits are key factors that help ensure that TEXT 2000 provides a foundation for meeting the State's document management needs for several years to come. They consist of

- Future compatibility
- Scalability
- Extensibility
- Future support for public access

Immediate Benefits

Improved Throughput and Turnaround: The TEXT 2000 solution will speed up the handling of legislative documents by eliminating non-essential manual tasks, automating others, and introducing work-flow management. This will result in better customer service from the legislative support agencies to legislators and the public. Benefits will be most obvious during busy legislative sessions when the volume of work is greatest and the State is working under tight time constraints.

Consider, for example, the impact upon the veto process. Wisconsin's governor enjoys line-item veto power over the budget, and approximately 1,300 pieces of text are vetoed each year. Before TEXT 2000, each of these vetoes meant that a piece of adhesive vellum must be cut to the same shape of the text and applied to the document before printing. This process alone took up to a week to complete. With TEXT 2000, vetoed text is marked electronically and the system highlights it automatically when the revised bill is printed.

Improved Quality: Capturing information at the source and eliminating re-keying , will improve the quality of critical State documents. TEXT 2000 's proofing and validation tools will also improve quality and result in less re-work time.

One example of this can be found in the documents prepared by the Chief Clerk's Office. Information which was keyed in multiple times to produce several documents is now keyed in once (see Figure 2), reducing the possibility of errors.

> *"This will result in better customer service from the legislative support agencies to legislators and the public."*

Improved Usability: Providing people with intuitive and straightforward tools is important in any business process. TEXT 2000 moves the State from a mainframe system with complex formatting codes to client/server with WYSIWYG and a menu-driven window structure.

More Flexibility: With TEXT 2000, the State will create new types of documents from existing information or documents, pull information from multiple locations, and publish information in multiple formats.

Longer-Term Benefits

Future Compatibility: The TEXT 2000 solution incorporates the latest in document-management technology. The products that make up TEXT 2000 are state-of-the-art. The customized off-the-shelf software solution allows for changing or updating individual pieces of the system over time, rather than replacing the entire system. The open system client/server approach also allows integration with other open systems.

Scalability: State agencies besides those designated to receive TEXT 2000 have already indicated interest in using it. A scalable architecture provides the State with flexibility to add additional agencies and users to the system. Additional users can be added incrementally in an easier distributed approach. Also, the scalable architecture allows individual pieces of hardware to be upgraded, giving the State more flexibility.

Extensibility: TEXT 2000 products have open APIs that are easy to integrate with other products and with custom routines. These products can be further customized as new requirements emerge, and new products can be integrated with the system as they become available.

Future Support for Public Access: Public access to public information through TEXT 2000 is a long-term goal. TEXT 2000 will provide a foundation that someday could allow the public dial-in access to statutes and other legislative information.

LESSONS LEARNED

Organizations planning large document management projects can learn a lot from the TEXT 2000 project. Described below are some critical issues that the State of Wisconsin faced, and the lessons that it drew from this experience.

SGML

The TEXT 2000 system had to handle highly structured documents with complex tagging requirements—a requirement that would typically suggest an SGML solution. However, this requirement had to be balanced against a requirement that the drafting tools be easy to use and intuitive to the end-user. One of the primary goals, in fact, was to get as many users as possible involved in text entry.

Interleaf was able to meet the State's formatting requirements through its support of templates. With Interleaf, users have the ability to tag text for use in document shredding, searching and composition. While providing an easy to understand user interface, Interleaf is also extensible and supports an open API, allowing it to be integrated easily into the TEXT 2000 solution.

Interleaf also offers an SGML product, but adopting it instead would require more knowledge about document structure, etc., than a typical TEXT 2000 user has. After weighing all the requirements, the project team decided that to make the system as intuitive and accessible as possible, SGML was not appropriate for TEXT 2000.

Criteria for Object Definition

Because the repository is object-oriented, the criteria for defining new classes to represent

documents had to be examined. Early on, the project team had to decide which documents should be stored as a single object class, which documents should be differentiated by an attribute value to identify type, and which documents should have their own object class.

The TEXT 2000 project used four criteria or guidelines to determine when a new object class should be created for a document. Some of these rules are hard and fast; others are ambiguous and may not always lead to the same answer. Any given project may have a different set of rules.

Creating too many classes or too few classes can make the system difficult and cumbersome to use. In general, new object classes should not be created unless one of the following rules is met:

Metadata: Is there specific metadata associated with the document? If so, a new class should be created for it. For legislative drafts, for example, it is important to know who requested the draft, who was responsible for drafting it, when it is needed, its bill number, etc. "Metadata" such as this is specific to legislative drafts, and, as such, it needs to be defined as part of a separate object class.

Behavior Modification: Does the standard behavior for a document need to be modified in any way? Certain documents need to be shred to the database when they are checked back into the system. Because the standard behavior for checking in this type of document is different, it should be defined as a separate object class. This allows the system to modify behavior appropriately.

Custom Extensions: Are any custom methods (programs) required that are specific to a type of document? Wisconsin's Statutes, for example, have specialized routines for proofing the content and checking for errors. Because this custom extension applied only to the Statutes, they should be defined as a separate object class This is the easiest way for the system to identify which documents to associate with the routines.

User Job Definitions/Implied Groupings: Some users are required to work with only one type of document; in other cases, users may associate certain documents as belonging to a special group or type. If so, creating a separate object class for these documents will tend to make working with the documents easier in terms of working with the documents, searching for documents and creating new documents.

Technical Integration Issues

The TEXT 2000 solution is built upon two different operating systems, two different processor types, and seven different software packages (with seven different programming languages):

- *Operating Systems:* DOS, UNIX
- *Languages:* C/C++, LISP, GAWK, C-Shell, Visual Basic, Cloverleaf, SQL
- *Integrated Products:* Documentum, Oracle, Interleaf, Smartleaf Database Publishing, Smartleaf Compare, Oracle Coauthor, Adobe Acrobat

The complexity of the solution's components presented a challenge to the development team, to integrate the products across the different platforms using different tools.

The most important part of the integration effort was creating a seamless interface between Interleaf and the Documentum Workspace. The requirement was that when a user opens a document for edit in the Documentum Workspace, the document is opened in Interleaf. This presented quite a challenge, since the Documentum Workspace resides on the workstation, and Interleaf runs on the Application Server.

"Creating too many classes or too few classes can make the system difficult and cumbersome to use."

This issue was tackled primarily because both products offered an open API, which allowed the development team to create the custom extensions required to integrate the two. Not only does this integration require the ability to check out a specific *document* to Interleaf, but also compound documents, or *books*. With TEXT 2000, users can check out and edit any piece of a book (or the entire book) and the portion that is checked out is given to Interleaf as a single unit.

Technology Assimilation

The TEXT 2000 user group consists primarily of legislative support staff and attorneys. Their computer literacy varies greatly; several attorneys, for example, did not even use a typewriter, much less a computer, before TEXT 2000. Although the user community as a whole was very excited about TEXT 2000 (as it offered huge improvements over the old system), some users worried that they would not be able to learn the new system, or feared the loss of expertise they had achieved under the old system.

These fears were dealt with by a proactive approach to help the State assimilate the new technology. This approach emphasized training and involving users in system planning and system support.

Training: For some users, TEXT 2000 will be the first Windows-based system they will have used, or the first time they have used a mouse. The concept of working with an applications toolkit such as TEXT 2000 is also new for most users. To lay the groundwork for these users, basic Windows and mouse tutorials (as well as games) were made available to help users learn Windows concepts in a non-threatening environment. Users could also practice their keyboard skills with on-line typing tutorials.

Self-paced tutorials were also prepared on TEXT 2000 applications. These tutorials allow users to experiment with these applications in a safe environment, *i.e.*, without worrying about damaging any files or data. Users found it easiest to work with these tutorials early in the project, when there was still plenty of time before the "real" system was implemented.

Involving users in systems planning: Getting users involved in the project and encouraging interest has been a great way to prepare them for TEXT 2000. Regular TEXT 2000 newsletters are published highlighting project news, questions and answers, and other information related to the system. Demos also allowed users to see the system first-hand. By seeing the system in action, users start to imagine what the system will be like in their own work setting.

Users are also involved in decisions related to all aspects of the project, including design, training, and rollout decisions. This increases user ownership over the project and excitement as they see their ideas implemented.

Involving users in system support: TEXT 2000 support will be provided by the State's "help desk," by its system administration group and through several "user experts". The user-expert concept was born out of the agencies' desire to have local experts who understood both the agencies' business processes and the TEXT 2000 system. The user experts should help agencies solve problems independently and quickly, rather than relying only on a central help desk. The user experts will be particularly suited to solving user problems arising from custom features developed under TEXT 2000 to address the needs of particular agencies or departments.

Organizational Impacts

When implemented, TEXT 2000 will cause several organizational impacts. These need to be anticipated and addressed to ensure a smooth implementation. For example:

"The requirement was that when a user opens a document for edit in the Documentum Workspace, the document is opened in Interleaf."

- Some job tasks will be automated or eliminated
- People may perform different tasks than they did before
- New skills will be required to use TEXT 2000
- Communication between work groups will change

These changes will help the State realize the benefits of TEXT 2000, but they can scare workers who do not yet understand how they will be personally impacted by such change.

Preparing the State's workforce for these changes involves more than just training. It has involved communication—helping users see the TEXT 2000 vision of the future, and how it is different from the way work is done today. For example, the vision includes much more active use of the system by lawyers than today. Some lawyers must decide whether to start using a computer for the first time ever or to continue drafting by hand. The lawyers have been actively encouraged to participate in demonstrations and tutorials, and more and more are now growing enthusiastic about the system.

Users also need to know details of the revised workflows, and what their future roles in them will be. For example, users were involved in reengineering the legislative drafting workflow, and once the new plan was decided upon, users had several months to review and comment on them before they were implemented. Receiving these new workflows early tempered user anxiety over the proposed changes (and undercut the rumor mill).

Proper preparation also means ensuring that the proper infrastructure is in place to support the revised workflows. The State of Wisconsin established channels of user support early, so that these support providers could themselves be trained early and then could stay more involved in the project than other users needed to be.

CONCLUSIONS

The TEXT 2000 project of the State of Wisconsin Legislature is an example of a large, mission-critical document management system that is being implemented as part of a broad process reengineering effort while older systems are still in place. Like most document management projects of this size, it is unique in many respects; still, many lessons that the State learned in this project can be applied to other corporate and governmental document-management projects.

In this case, custom integration of off-the-shelf software components resulted in a system which is easy to maintain, and which meets the particular needs of several different legislative departments or agencies. Selecting extensible products with robust applications programming interfaces (API's) aided both in the integration of system components and in the extension of these components to address custom requirements of various agencies.

Critical to the success of this project was involving end-users—even those who had not used computers before—in the design and support of the new system. Involving users early in the design and planning process allayed their fears about the potential impact of TEXT 2000 on their jobs or livelihood. Tutorials and demonstrations of the new system provided a similar function for those users who chose not to become directly involved in the systems-planning process.

Ansdersen Consulting: Michael Castle, Michael Polelle, Michelle Miller

INDUSTRY NEWS

With this issue we are expanding our report to cover news items we think are most relevant and interesting to our readers. We will also use this section to comment on how industry news and events (including conferences and trade shows) are affecting the document management and document computing marketplace. This section will no doubt grow as we proceed. If you have news, keep us posted, and if you have comments on what you would like to see us cover, let us know.

INTERLEAF ANNOUNCES NEW DOCUMENT MANAGEMENT PACKAGE

Interleaf will be announcing Intellecte at the Documation '94 conference on February 22. Intellecte is described as a complete document management system that will allow corporations to quickly build a repository of document information with WYSIWYG "point and click" access for 100 users.

Intellecte is an integrated packaged solution including a number of Interleaf's products (such as RDM and WorldView), as well as the consulting services required to implement support for five chosen document types. The entire system is designed to be up and running in 12 weeks.

By creating a well defined package, Interleaf hopes to dramatically ease the process of implementing a document management system, and to keep the cost low. Standard pricing will be provided at the announcement.

SGML OPEN CONSORTIUM APPOINTS EXECUTIVE DIRECTOR, NAMES TREASURER, & ORGANIZES FOR CONFORMANCE TESTING

Mary Laplante has joined the consortium as its executive director. Laplante, a well-known marketing consultant, assumed the director's role on January 1, 1994. In addition to Laplante's hiring, the Directors announced that Robin Tomlin, Senior Manager at Intergraph Corporation, has joined the SGML Open Board as the consortium's secretary/treasurer.

SGML conformance testing has been something many corporate and government users have longed for. The Computer Systems Lab at NIST (National Institute for Standards & Technology) has been funded and is working on guidelines for certifying conformance with federal government SGML requirements. However the NIST proposed policy and procedures for testing and certification has raised a number of concerns among both SGML vendors and users.

To ensure that any NIST guidelines meet the needs of the SGML community, SGML Open has joined up with the Graphic Communications Association and the International SGML Users Group to review the NIST work and make suggestions to alleviate industry concerns. For conformance testing to succeed NIST needs the backing of private industry.

Oracle Steps Up Move into Document Management

The week after we go to press with this issue Oracle will be discussing its plans for text and multimedia support at its Media Server rollout in Los Angeles. Oracle plans to integrate its text retrieval (ConText) product with their electronic viewing (Oracle Book), and office (Oracle Office) products.

At this point it isn't clear just how all these are going to be tied together with each other and the Media Server, but it sure looks like Oracle is stepping up its move toward the document management market.

Image Business Systems Has New CEO and Scandinavian Distributor

Image Business Systems has appointed John C. Daily as its new president and CEO. Previous CEO Jay Goldberg will continue as Chairman of the Board and remain involved in strategic planning.

Separately IBS announced that W.S.A. Scandinavia AB has signed a three year contract to sell the Imagesystem document management & workflow product in Norway, Denmark, Sweden and Finland.

Dataware To Acquire BRS

Dataware Technologies has announced an agreement to acquire the BRS Software Product segment of InfoPro Technologies. The acquisition continues Dataware's rapid growth over the past couple of years, and comes about a half a year after their successful IPO. While both existing product lines will continue to be sold and supported, it will be interesting to see what new products will come from the combined development groups.

In a separate announcement, BRS Software Products announced that it is shipping its new BRS/Word Plus product, which provides word processing conversion capability to facilitate input to its document management system.

FileNet Reports Earnings

FileNet reported estimated earnings of $45.9 million for its fourth quarter ending in January. This represents an increase of 44% over last years fourth quarter ($31.8 million). Estimated revenues for the year are $158.8 million with net income of $7.8 million.

FileNet attributes the growth to an increase in software sales to business partners (HP & IBM), continuing their transition to a software oriented company.

Open Text & Booz Allen Announce Alliance

Open Text and Booz Allen & Hamilton announced an agreement that calls for the incorporation of Open Text's text retrieval engine (Pat) into Booz Allen's information integration software (Minerva). The companies expect the alliance to generate over $1 million to Open Text.

CALENDAR OF EVENTS

Below is a selection of key events covering open information and document system issues. There are many other conferences and shows covering related topics. We will attempt to keep this list to those events that focus on areas most directly related to the areas covered in our report.

Documation '94. February 21-25, 1994, Los Angeles CA. The annual international event for document management applications and document computing. Call (703) 519-8160 or (617) 643-8855, Fax (703) 548-2867 or (617) 648-0678.

Intermedia. March 1-3, 1994. San Jose, CA. Multimedia and CD-ROM. Conference and Exhibition. Call (203) 352-8240, Fax (214) 245-8700.

Seybold Seminars '94. March 22-25, 1994. Boston, MA. The annual gathering of the computer publishing elite. Conference and Exhibition. Call (310) 457-8500, Fax (310) 457-8510.

OnLine Publishing '94. April 10-13, New York, NY. GCA conference on online publishing issues. Call (703) 519-8160, Fax (703) 548-2867.

AIIM. April 18-21, 1994, New York, NY. AIIM's annual show and conference focusing on imaging and storage and retrieval. Call (301) 587-8202.

Pen & Portable Computing. May 2-5, 1994, Boston, MA. Well, Documents need to be portable don't they? Sponsored by Boston University. Call (800) 733-3593, ext. 255, FAX (508) 649-2162.

EDD '94. May 10-12, 1994, Somerset, NJ. Bellcore's forum for discussion of issues relating to the exchange of technical information in electronic form. Call (201)829-4135, Fax (201)829-5883.

SGML Europe. May 15-19, 1994, Montreux, Switzerland. The European counterpart to the SGML '93 conference in the U.S. Call (703) 519-8160, Fax (703) 548-2867.

AIA Automated Technical Data Symposium & Exhibition. May 16-18, St. Louis, MO. The 9th biennial gathering of the Aerospace Industries Association group focused on managing technical data. This year's theme: Interactive Electronic Environments. Call (202) 371-8435, Fax (202) 371-8470.

Vision '97 - Vendor Interaction Symposium. April 27-29, 1994. Phoenixs, AZ. Xplor's printing vendor get together to discuss the future. Call (214) 867-4228, Fax (214) 964-3361.

Seybold Paris. June 8-10, 1994. Paris, France. Seybold's main European event. Conference and Exhibition. Call +44 (0)323 410561 , Fax +44 (0)323 410279.

Infobase '94. June 28-30, 1994. Salt Lake City, UT. Folio User Conference. Call (801) 344-3671, or (801) 344 3672.

International Conference on HyTime. July 24-27, 1994, Vancouver, BC Canada. New conference exploring applications of the ISO standard. Call (703) 519-8160, Fax (703) 548-2867.

CALS Europe '94. September 14-16, 1994, Paris, France. The annual pan-European conference on CALS technology and applications. Call (703) 578-0301 or +49 30 882 6656, Fax (703) 578-3386 or +49 30 883 8811.

TOPICS COVERED IN PREVIOUS ISSUES

Vol. 1, No. 1.

What The Report Will Cover & Why — An Introduction To "Open Document Systems", And A Description Of The Report's Objectives.

Imaging, Document & Information Management Systems — What's The Difference, And How Do You Know What You Need?

Vol. 1, No. 2.

SGML Open — Why SGML And Why A Consortium?

Document Query Languages — Why Is It So Hard To Ask A Simple Question?

Vol. 1, No. 3.

Document Management & Databases — What's The Relationship?

Vol. 1, No. 4.

Electronic Delivery — What Are The Implementation Issues For Corporate Applications?

Vol. 1, No. 5.

Multimedia Rights & Wrongs — What IS Managers Should Know About Copyrights In The Age Of Multimedia.

Vol. 1, No. 6.

Document-Centered Interfaces & Object-Oriented Programming — How Will They Affect You?

TOPICS TO BE COVERED IN FUTURE ISSUES

The subjects listed below are some of the areas we will be covering, in no particular order. If you have an opinion about which topics you would like to see added or covered sooner rather than later, let us know.

Office Workflow Systems — Can They Handle Strategic Information, Or Are They For Casual Or *Ad Hoc* Use Only?

SGML & Presentation Interchange — What Standards Are Available Or Appropriate? (DSSSL, OS/FOSI, HyTime, ODA, etc.)

Authoring Systems — Do You Need Different Kinds For Different Media?

ISO 9000 — What Kind Of Document Management System Do You Need To Meet This Quality System Standard?

The Airframe And Airline Industry's Strategy For Sharing Product Information — What Can You Learn From It?

New Drug Applications — What Document System Strategies Make Sense For The Pharmaceutical Industry?

Object & Relational Databases — Which Approach Is More Suited To Your Document Systems Needs?

Compound Document Architectures — Why Do We Need Them? Who Will Define Them? Will They Do What We Expect?

SGML Versus ODA — How Do They Differ? Is There A Reason To Have Both? What Can They Do? Which Approach Is Right For The Future?

Order Form

o Please start my subscription to: The Gilbane Report on Open Information & Document Systems (6 issues). Back issues available for $45. each.

U.S.A.: $225   Canada: $232   Foreign: $242

Additional copies and site licenses are available at reduced rates. Call for information.

Please send me additional   o Consulting Services   o Special Reports
information on: o The Annual Documation Conference & Exposition
                                o Document Managment & Electronic Delivery Seminars

o My check for $ _____ is enclosed   o Please bill me
o Please charge my credit card   o MasterCard   o Visa   o American Express

Name as it appears on card _____   Number _____

Signature _____   Expiration date_____

Checks from Canada and elsewhere outside the U.S. should be made payable in U.S. dollars. Funds may be transferred directly to our bank: Baybank Boston NA, 175 Federal Street, Boston MA 02110, S.W. code BAYBUS33, into the account of Publishing Technology Management, Inc., number 1444-89-63. Please be sure to identify the name of the subscriber and the nature of the order if funds are transferred bank-to-bank.

Name _____   Title_____

Company _____   Department_____

Address_____

City _____   State _____   Zip_____   Country _____

Telephone _____   Fax _____   Email _____

**Mail or fax this form to: Publishing Technology Management, Inc., 46 Lewis Avenue, Arlington MA 02174-3206
Fax: (617) 648-0678 • To order by phone call: (617) 643-8855**

HOW TO FIND OUT MORE ABOUT COMPANIES MENTIONED IN THIS ISSUE

Adobe Systems. Inc.
1625 Charlestown Road
Mountain View, CA 94043
(415) 961-4400

Andersen Consulting
777 East Wisconsin Avenue
Milwaukee, WI 53202
(414) 283-3000

Andersen Consulting
801 Second Avenue, Ste 900
Seattle, WA 98104-1509

DataBase Publishing Software, Inc.
400 Cummings Park, Ste 5300
Woburn, MA 01801
(617) 938-0018

Data Conversion Labs
184-13 Horace Harding Expway
Fresh Meadows, NY 11365

Dataware Technologies
222 Third Street, Ste 3300
Cambridge, MA 02142
(617) 621-0820

Documentum
5724 West Las Positas Blvd.
Pleasanton, CA 94588
(510) 460-4115

FileNet
3565 Harbor Blvd.
Costa Mesa, CA 92626
(714) 966-3400

Image Business Systems
417 Fifth Ave.
New York, NY 10016
(212) 696-2500

Interleaf
Prospect Place, 9 Hillside Ave.
Waltham, MA 02154
(617) 290-0710

Microsoft Corporation
One Microsoft Way
Redmond, WA
(206) 882-8080

Open Text Corporation
Suite 550,
180 King Street South
Waterloo, ON N2J 1P8
Canada
(519) 571-7111

Oracle Corporation
500 Oracle Parkway
Redwood Shores, CA 94065
(415) 506-7000
(416) 239-4801

Sun Microsystems
2550 Garcia Avenue
Mountain View, CA 94043
(415) 960-1300

Verity
1550 Plymouth
Mountain View, CA 94043
(415) 960-7600