THE GILBANE REPORT

on Open Information & Document Systems

DOCUMENT MANAGEMENT INDUSTRY UPDATE — DOCUMATION '94 & OTHER SPRING ACTIVITY

In this issue we depart from our usual format. Instead of analyzing a particular issue or using a case study to illustrate trends in the document management industry, we look closely at some exciting events at the Documation '94 conference and exposition. The only event covering a wide spectrum of document management technology and business solutions, — not just image storage and retrieval and/or printing systems. Documation this year was where the most important announcements and activity took place. Throughout this issue we also present important relevant announcements from other Spring conferences, (including Seybold, AIIM and Interop) especially as they relate to the document management market.

COMING IN AUGUST

Our next issue will take on one of the most difficult (and most controversial!) aspects of document interchange — that of exchanging formatting and presentation information. This problem is variously viewed as: "unsolvable", "trivial", or "not even a problem". Experts disagree, but without a doubt many people want a solution. Companies need to understand the underlying issues before settling on a strategy for managing documents and the information they contain.

WE'VE MOVED!

PTM and The Gilbane Report have moved into new offices in Harvard Square. See the back cover for our new address.

CONTENTS

DOCUMENT MANAGEMENT INDUSTRY UPDATE —

DOCUMATION '94 & OTHER SPRING EVENTS

EXECUTIVE SUMMARY

Strategic Overview

• The "Document Management" market is by all accounts in the early stages of a rapid rise. Various market researchers forecast an annual growth rate of 20-25%.

• As a market, however, it is still poorly defined. The hundreds of different types of document management products, from a wide variety of vendors, combine to form a confusing picture.

• The need for document management is beyond dispute. Progress in technology and some critically important market trends require careful monitoring and evaluation by IS and IT managers.

Document Management Trends — Documation And Other Spring Events

• The success of Documation, the growing number of trade shows that address some aspect of document management, and the sheer number of document management related announcements, reinforce market research projections.

• Signs of the industry's growing pains appear as different types of vendors try to navigate around the expanding territory, and position themselves for success. Alliances are picking up speed. Further consolidation activity will be something to keep an eye on.

On the positive side, such growth is teaching vendors that they need to work together to create and promote, interoperability standards that businesses need to make effective use of various technologies.

• Document management functionality is finally beginning to be seen as a mainstream business need. As a result it is emerging as a fundamental requirement of enterprise IT architectures.

• Technology developments involving document computing are accelerating and will eventually surpass those limited to data processing.

Conclusions

• Users should think of document management as an integral part of their information management strategies, not as just a niche application. (There will, however, continue to be a need for niche applications). Managers must be especially diligent to make sure they are employing the right document management techniques and approaches for the business problems they are trying to solve.

• Finding the right document management approach requires including document reengineering as part of any business process reeningeering.

- Users should be aggressive in demanding application interoperability, and should push suppliers for standards.

- Vendors must realize that the market will not only grow, but change dramatically as a result of the growth. Their products must interoperate with many different kinds of software applications, including those from direct competitors, from indirect competitors (for example, those with an entirely different approach to solving the same problem), and mainstream office applications.

- Platform vendors, and to some extent, database vendors, need to be sure to provide the kinds of infrastructure support that sophisticated document management applications will require.

- Integrators, large and small, should recognize the vast opportunity for helping companies cope with the bewildering array of document management products and technology available. The amount of integration required, especially for "enterprise" solutions, ensures a healthy market for outsourcing, as well as for project based integration

STRATEGIC OVERVIEW

Market Growth

By any standards, the Document Management market is growing extremely rapidly. What is usually called the document imaging or document image management segment is estimated by market researchers to be between $1.5 to $2 billion[1] , and such forecasts usually exclude compound document management and integrated document management. Sales of certain other components of document management systems, such as optical character recognition and full-text retrieval, may or may not be included, depending on how each researcher or study defines document management. Expected growth rates in all the market studies we have seen recently are high; ranging from 15 to 25 percent per year for the next 3-5 years.

Market Confusion

Many different kinds of "document management" systems and products are on the market today — literally hundreds. This is good news — it implies that there are enough choices that one of them should meet your needs. The bad news, however, is that few of us have either the time or inclination to sift through more than a handful of products.

One of the reasons that so many choices exist is that the term "document management" itself is used to describe so many vastly different types of products and activities: everything from a basic word search capability in a word processor, to a set of UNIX directory naming conventions, to a scanning subsystem to a database cataloger to a printer to an electronic viewer. This mess is further compounded by "workflow" and "groupware" products, which are often sold, sometimes appropriately and sometimes not, to solve document management problems. This also explains why you see "document management" products on display too at so many different kinds of trade shows.

The forecasted market growth also ensures that more, (and larger,) companies will be entering the market. The supplier community is under severe pressure to deliver integrated solutions involving multiple products. These two forces will fuel the formation of new kinds of alliances, and guarantee a certain amount of consolidation as well.

---

[1] Take all estimates of the size of the document management market with a grain of salt. We have not attempted a serious statistical analysis or verification, but are merely summarizing what you might find by reading the different studies available today (which are not in agreement).

Technology Developments

The pace of development in all areas of information technology is sometimes dizzying, and this is no less true with document systems technology. However the next 12-24 months will bring some changes in computing operating environments that will provide a new foundation for document computing. This new infrastructure will profoundly enhance our ability to manage documents and the information they contain. (See *Vol. 1, Num. 6*, for an analysis of some of these issues).

Most businesses drastically underestimate how soon electronic delivery and multimedia documents will become prevalent — and indispensable — throughout their organizations. Already, voice annotation has become a standard feature of mainstream word processors — you get it at no extra cost. As voice mail becomes more integrated with E-mail, it won't take long for memos and reports shared on your LAN-based E-mail to include voice notes such as "Hector, click on the voice note icon to hear the message Sarah left me about the Glitsch account".

Challenge

All this activity increases the risk that IS managers will implement document management solutions that are inappropriate to their company's needs.

Savvy managers will first work to narrow down the scope of their searches — both for products and vendors who understand their business's needs.

DOCUMATION '94

This was Documation's first year; still, there was too much going on for one person to begin to cover it all. And although many of our subscribers attended the conference, we have received numerous requests from others to provide some coverage.

As the conference chair and cosponsor of the event, your editor was unable to visit and report on every session. Matt Shanahan, and Mary LaPlante, each of whom were track chairs of the conference, have kindly agreed to contribute their thoughts on important trends and milestones at the conference, supplementing our own report on what happened in the general sessions.

Even with Matt and Mary's help, our summary only begins to give you a feel for all that went on at Documation. We did not attempt to cover every session, and, in any case, an overview is no substitute for the real-time dialog and interactivity that takes place at such conferences and on the show floor.

Background[2]

Co-sponsored by PTM, The Gilbane Report, the Graphic Communication Association (GCA) and the Graphic Communication Association Research Institute (GCARI), Documation '94 attracted 573 conference attendees; over 2000 others attended the accompanying exposition. Attendees and speakers came from 12 countries, and the 45 members of the press included representatives from Europe and Asia.

The "average" profile of attendees mapped pretty closely with our readership — that is, mainly corporate IS, IT, and line managers from vertical industries across the globe. Not surprisingly, the industries most strongly represented were those under the most pressure to get their documents under control due to legal, or regulatory restrictions, industry

> "*Most businesses drastically underestimate how soon electronic delivery and multi-media documents will become prevalent — and indispensable — throughout their organizations.*"

---

[2] The proceedings, which include many (not all) of the presentations are still available from the GCA at (703) 519-8160.

standards, or competitive pressure to improve service or to create customized information products.

The Document Management Industry

The opening remarks, provided by your editor, outlined the state of the document management industry in terms of: the business problem, the technology, and the evolving vendor landscape.

The Challenge and the Business Problem. Many different numbers have been cited as representing the amount of corporate resources swallowed-up in the management of documents. Whether the proportion of corporate information in document form, (as opposed to database form,) is 80% or 95%, still the vast majority of corporate information is in documents, whether paper or digital. The percentage of corporate revenues required to manage this document information is typically estimated at somewhere between 5 and 15 percent — clearly an amount worthy of serious scrutiny by corporate financial officers.

Most businesses today, however, have done little, if anything, yet to manage documents. Those that do often manage only a small percentage of their documents; (perhaps one or two isolated applications hidden away in one department or another). As we often note, much of this document information is strategic, — that is, it is central to the company's business. Critical information such as product documentation, engineering drawings, financial records, customer profiles, or regulatory documents are strategic assets that are not often enough leveraged to increase business opportunities

Information is becoming a larger part of products, especially complex products that require lots of documents to support their design, manufacture, delivery and maintenance. The cost of products increases proportionally, so without ways to cut down the increased costs of creating, managing and delivering document information, companies have to raise their prices and be less competitive.

Given all this, a corporate information management strategy that doesn't include at its core a document strategy is inadequate, to say the least.

Document Management Technology. Technology is already available to manage documents in efficient ways. Our challenge today is to reengineer our business processes to use this technology effectively, to choose among the hundreds of tools and to integrate them into successful solutions. We are, however, in the early stages of the next frontier of information technology solutions — we have solved the data management problem (more or less), now it's time to solve the document management problem. Document-oriented computing is on the horizon, and it promises to drastically improve our ability to manage information in documents without first having to extract it and force-fit it into a database.

The Evolving Solutions Landscape. Vendors from many different areas are getting into the document management market. There are basically three categories of product providers with overlapping and converging capabilities for managing documents (See Figure 1). Document imaging vendors are continually adding workflow, security, and distribution features. Publishing system vendors have added the use of databases and workflow tools to their paper and electronic delivery offerings. Database providers are providing increasing capability to manage the complex information found in documents that are in turn used by application developers. This convergence will continue and will guarantee some interesting alliances and positioning.
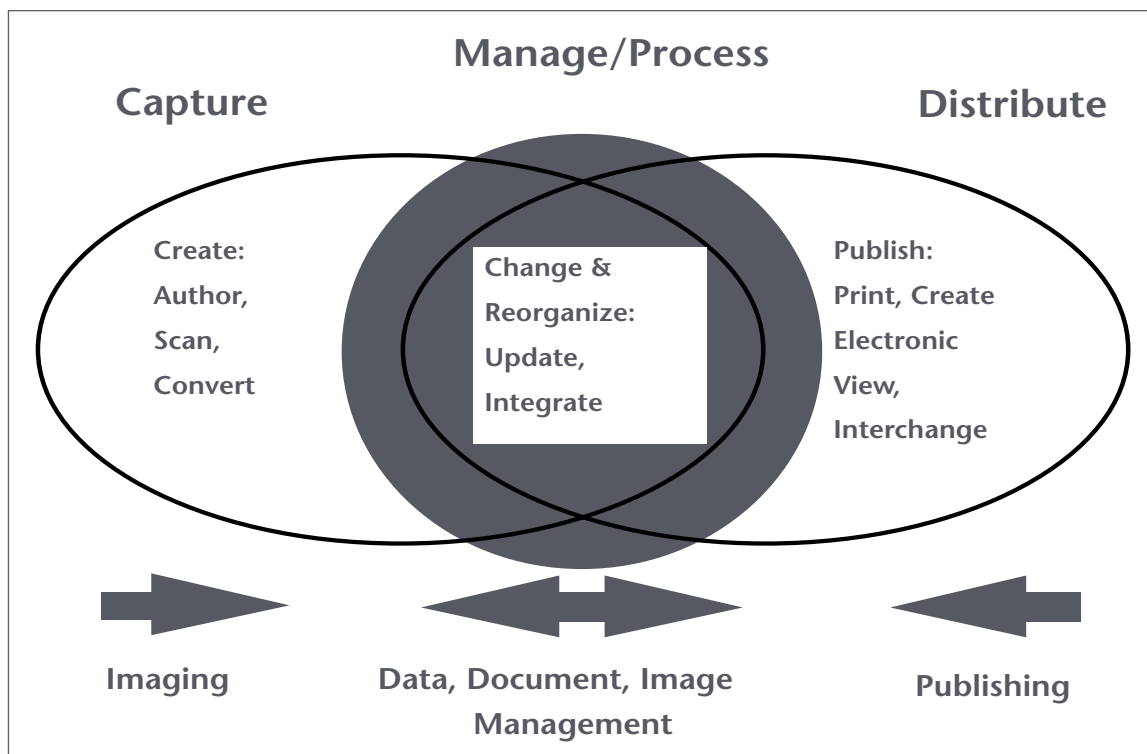
" … a corporate information management strategy that doesn't include at its core a document strategy is inadequate, to say the least."

Manage/Process

Capture

Distribute

Create:
Author,
Scan,
Convert

**Change &
Reorganize:**
Update,
Integrate

Publish:
Print, Create
Electronic
View,
Interchange

Imaging

**Data, Document, Image
Management**

Publishing

The business challenge, the availabale and coming technology, and the maturing supplier landscape combine to ensure that document system technology will drive information management strategies for the next few years.

The opening-session speakers were asked to describe: their vision of document management, and how they saw themselves working with the other types of suppliers in the marketplace.

Deborah Triant from Adobe spoke of Adobe's experience in adopting their own electronic delivery product, Acrobat, throughout the company. Describing the many benefits they gained, Triant added a word of caution. Her lessons-learned message was that going electronic is great, but it doesn't take away the need to have management systems in place to deal with all the documents.

Deborah also made the announcement of the Shamrock consortium on behalf of the member companies of which Adobe is one. (See below for a discussion on Shamrock).

Mark Ruport of Interleaf argued that document management is not simply about managing documents, but about managing the information in them — the information that makes the documents useful and valuable. He distinguished among three types of data: *transaction-based*, which is managed by database managers; *office data*, managed by personal productivity tools and file management-based document management systems; and *strategic data* that drives critical business processes that needs to be managed by document management systems that can assemble all the critical information in an enterprise and manage it.

Paperlessness, according to Ruport, is not a revolution, it's just an upgrade. If all you have done is create electronic versions of fixed documents, you haven't reengineered your business processes to make the most effective use of your information.

Mark ended his presentation with the announcement of their (then new) Intellecte document management product (See below for more information on products announcements).

"Paperless-

ness, accord-

ing to Ruport,

is not a revolu-

tion, it's just

an upgrade."

Larry Stevens spoke about Oracle's interest in the text and multimedia market. Just a week earlier Oracle had announced its Media Server and Text Server, and while Larry did not make any further product announcements, he couldn't resist giving us a quick demo of natural-language text-retrieval technology he just happened to have on his laptop. Recently announced as ConText, this technology is designed to be integrated with other Oracle products including the electronic viewer, Oracle Book.

FileNet's **Jordan Libit** spoke about the critical added value that workflow management adds to document and image management systems. He described a number of different applications to illustrate how document image management needs to be integrated with the information flow of business processes in flexible ways to gain real benefits.

Peter Lamb from Andersen Consulting provided a thorough overview of the document management market and the critical issues from the perspective of a consultant and integrator. In his view, "document and object management" is the focal point of convergence for many different information management activities: imaging, groupware, data processing, engineering document management, technical publications and word processing. The driving forces behind all these areas are: communication, the need for time compres-
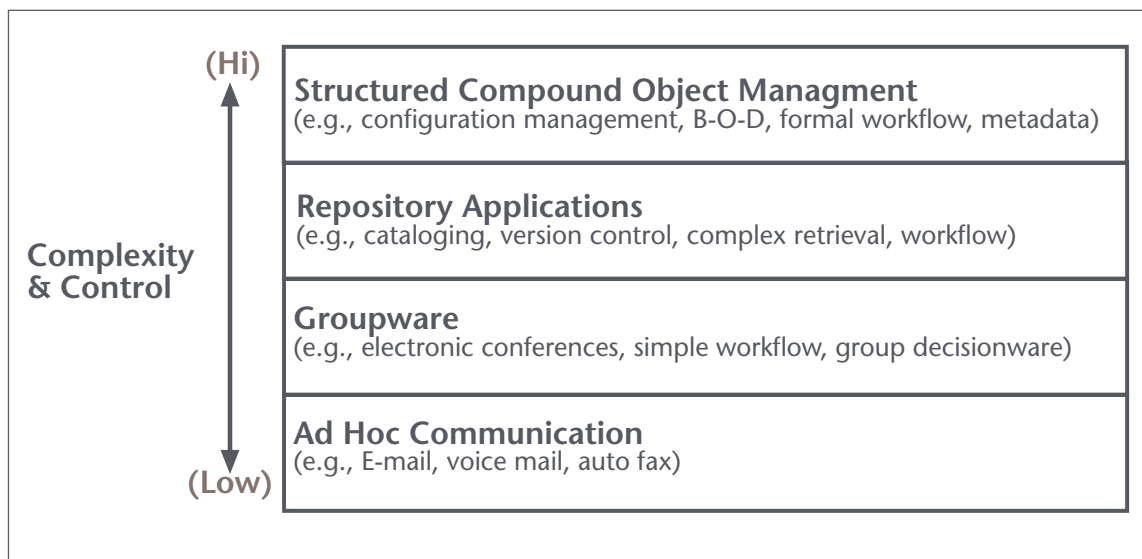

Figure 2

Peter's 4-Level

Framework

sion in all business activities (*e.g.*, time to market, decision-making, customer service), and effective use of knowledge as a competitive weapon in the global market. Document management is a key component for: communication, process reengineering, change management, knowledge incorporation, and decision support/employee enablement. Lamb categorized document management applications into "Peter's 4-level Framework" (Figure 2), and discussed a half dozen cases illustrating different types of document management business problems. The five big issues today, according to Pete, are:

1. Format and interchange standards[3]

2. The role and extent of OODBs

3. Integration with "legacy" environments

4. Refocusing from transactions and documents to decisions and events, and

5. Organizational impacts.

---

[3] See our next issue for our analysis of this area.

Technology Trends — Document Computing

On Wednesday the general session was divided into two sections. One covered new technologies being developed to enhance document computing and document management. The other presented senior managers from large corporations who described their own document management needs.

Your editor opened the technology session by describing three components of current document management systems, each of which presage future developments. *Objects*— whether in terms of object-oriented databases, object-oriented programming, or multimedia document component "information objects"— play a big role in making systems more flexible and capable of dealing with complexity. Building an *architecture* to manage and share distributed objects, and to link and assemble them into document form are requirements of many enterprise-wide document management solutions. Finally, the *document metaphor* is increasingly seen as the most effective and friendly way to interface not only with document management systems, but with information in general.

Today, these capabilities are built either at the application level, or as "middleware"." For many reasons (*e.g.*, application interoperability, performance, and ease of application development), it would help instead to have support for these capabilities at the operating environment level.

Previous attempts at compound document architectures to provide such an environment have failed. But this is clearly something we need, and eventually will get.

Whoever defines and builds such an architecture will be in a powerful position to dominate the IT market. We can expect fierce battles among the platform and architecture vendors to control this architecture[4]. The two leading candidates today are Microsoft's OLE, and the Component Integration Lab consortium's OpenDoc (based on Apple technology).

Larry Tessler from Apple described the "Information Tidal Wave" (his alternative to "superhighway") coming with the growth of electronic multimedia documents, and with the rapid building of electronic document repositories. IS managers will face severe new problems arising from the need to manage these repositories. Larry positioned OpenDoc as a core technology for supporting the management and assembly of these new kinds of documents.

Microsoft's **Tony William**s focused on user requirements for a compound document architecture. Compound documents should be thought of as "compound views" of information, and documents are just one form of information, and thus need to be handled as part of an information architecture. Information architectures in turn need to be able to manage many different types of multimedia data for both document and data applications.

A standard "containment model" is needed, Williams said, to allow applications to share and organize information objects. Previous attempts at standard compound document architectures, *e.g.*, ODA (Office, or Open Document Architecture) failed because they attempted to define a too restrictive representation. Such systems also need to handle *ad hoc* information (for example, that created with a personal information manager) as well as structured documents.

Tony emphasized the need to protect both user investments in information and developer investments in applications. While a compound document architecture environment is a

"Previous attempts at compound document architectures to provide such an environment have failed."

---

[4] For a fascinating discussion of architecture battles in the computer industry, see 'Computer Wars', by Charles Ferguson and Charles Morris, published by Times Books, 1993.

requirement of any new operating environment, there must be an evolutionary path provided — a compound document architecture that forces a radical change too quickly will not gain acceptance. Tony positioned OLE as the technology that meets these requirements.

When asked, both Tony and Larry Tessler claimed that OpenDoc and OLE should work together and described generally — each in terms of the architecture they were promoting — how that could happen. However, this is definitely an area where there needs to be continued and aggressive vigilance on the part of corporate users to ensure that operating environment interoperability results. It would certainly not be wise — at least not yet — to assume that one of these approaches will become dominant.

Sun's **Bruce 'Tog' Tognazzini** delighted the audience with his presentation on the user-interface of the future. He started off with a demo of a user interface we will always have available. Rather than show us a slide illustrating the exponential growth of information and our losing battle to deal with it, he became a human graph by crouching down at one end of the room and running to the other end, all the time straightening-up and ending up in mid-air!

'Tog' showed a video of an office environment of the future where workers interfaced with real time multimedia information and with live video views of each other. The video gave new meaning to the term "desktop interface," since there was no distinction (from a user-interface point of view) between the desk and the computer screen. Live documents lived on the desk surface along with digital video conferenceing windows, both of which could be physically moved around at will.

What Senior IS & IT Managers Want

In another session, we invited senior managers from different industries to share their document management requirements with us, particularly for the benefit of the vendors in the audience. Panelists included: **Ed Jowdy** from Aetna Life & Casualty; **Fred Mitchell** from Boeing; **Charles Popper** from Merck & Company; and **Don Hedeen** from General Motors.

One goal was to determine what common requirements might exist among industries where document management is receiving a lot of attention. Another was to find out which technologies and services were at the top of the wish-list for such large purchasers consumers of information technology products and services.

Our choice of industries was not meant to be exclusive, only representative. Many companies in many different industries are trying to solve the same fundamental problems, but often the lack of a common vocabulary and dialogue across industries obscures common requirements.

All our panelists said they were motivated by customer service, since document and information delivery was recognized as a major problem for both internal corporate and external customers. No one doubted that document management systems could reduce costs, but they said they faced more pressing, service-oriented challenges as a result of rapid changes in the computing environment, and the vendor environment they face.

These challenges manifest themselves in two ways. First, they wanted suppliers to make it easier for them to integrate and manage different networks. It is still far too difficult to connect many different platforms in mixed PC, UNIX and legacy mainframe environments, especially when all are in a constant state of change.

Second, they strongly advised vendors to work together to make application interoperability[5] a reality, not just a buzzword. Some of the panelists were members of the Shamrock consortium (see below), but the others, too, agreed that this area was critical; some were already working with suppliers to deal with this problem.

The panelists made it clear that their first priority was making sure all their systems could communicate. They reassured applications vendors that the reason their solutions were not getting the attention they deserved was because there were still network infrastructure problems that had to be solved.

There were different points of view concerning standards. Everyone agreed that standards were critically important, but, what they meant by this was not always the same. Also, not everyone believed standards were (or would be) sufficient. For some SGML was essential; others were more concerned with specific problems of interoperability and standard APIs. Some, while conceding that standards are necessary, seemed to do so reluctantly; one sensed a nostalgia for an era when companies could skirt the issue by standardizing on a particular vendor.

Other differences were noted in the level of control that companies have over the documents they manage, and in the timeframes during which documents must be delivered. These clearly are two reasons why different companies take different approaches to document management.

Perhaps the most significant (or at least, the most welcome) message that the panel delivered to vendors was that they all viewed document management as a critical component of an information management infrastructure. The clear implication is that document management products will find a wider market once companies like these solve their more immediate connectivity and networking problems.

Electronic Delivery

Our general session on Thursday was designed to set the stage for the day's discussion of document delivery. In many ways the delivery of documents is the most important part of the document management process —this is where the customer first sees it. The challenge is to deliver a document that is accurate and is delivered when and how it is required.

Delivering paper documents, using both on-demand, and more traditional printing technology will continue to be a requirement. Paper after all, is the only widely accepted standard form for document delivery, and is extremely portable.

Electronic delivery however provides too many benefits for companies to ignore: It can reduce costs, increase delivery speed, improve information access, and enrich (documents with multimedia) information.

Industry today is in a transition to a world where information must be provided in whatever form the customer wants it. This means not only providing information in either electronic or paper form or both, but organizing and packaging documents in customized ways for different business partners. Whether this is a current requirement or not, managers need to have a strategy in place for the inevitable demand to come.

Our speakers were asked to provide some guidance to managers looking to develop a modern effective document delivery strategy.

Ed Heresniak had just left McGraw-Hill and did not officially represent them. Ed's remarks reflected his views as a purchaser and user of technology, and revolved around three

---

5. 'Interoperability' can be defined as 'the ability for applications to work together, and share information to achieve a solution to a business problem.'

observations: (1) if you can conceive it you can do it, (2) we need to think in "new templates", and (3) technology is still too difficult for most people.

The problem in delivering information, and of managing information in general, is not technology. The technical community is capable of developing solutions for any business problem we can conceive. However, we need to think in terms appropriate to the capabilities that new technology provides. Thinking in terms of static linear document form for example, is not the right way to learn in a dynamic interactive information environment. We need to immerse ourselves in information and learn new ways to extract the information we need.

Technology, is still too hard for most people. Ed pointed to the complexity of creating material for multimedia presentations, (His own presentation was "unimedia" — without any visual support) and he echoed Ed Jowdy's call for vendors to work closer on standards and to make products easier to use.

R.R. Donnelley's **Ron Brumback** provided an example-rich presentation that reinforced Heresniak's basic messages. The problem we face in delivering vast amounts of documents is not technology. The problem is that we need to change our use of the documents, the business processes, and people's roles. Businesses in general are ineffective in their use of the document information they have; they need a new information-handling infrastructure.

After giving three examples of businesses not adequately mining the information they had in documents, and not knowing how, or even if, the documents were being used by their customers. Ron listed three critical characteristics to watch for in creating more effective documents and for making sure they are used in useful and productive ways. The first is "sufficiency"; documents must provide the right amount of information, not too much but enough to do the intended job. Secondly, documents must be "available"; they must be delivered to the right people at the right time and place. Finally, documents must be "active"; they must in the form in which they can help the customer most. Documents may need to be in either paper or electronic form, or they may need to be customized, configured and presented to the customer in many different ways.

Dennis Andrews from XSoft had a simple straightforward message: "Keep your options open". The traditional approach to document distribution is to design and create multiple documents and then marry the source information to the document design. This causes problems as you grow and need to deliver new types of documents. This approach is no longer feasible given the cost and management challenges associated with re-creating new documents. The answer is to store the information once in a way that allows you to distribute many times, and in a variety of ways.

A document, after all, is simply a view of information objects. These objects should be managed in a database, and extracted for document creation, as necessary. Further, how information is presented — how documents look are designed — is critically important. Companies need to be able to format documents appropriately for the media and audience.

Other Important Announcements

Among the many announcements made at Documation, were a few whose importance went beyond particular products, and signified larger trends in the industry.[6]

> *"The problem is that we need to change our use of the documents, the business processes, and people's roles."*

---

6. (For more information on particular products announced at Documation, see the Seybold Report on Publishing Systems, vol. 23, no. 14, and OII Spectrum, vol. 1 no. 7).

Shamrock

One of the most important announcements at Documation was the formation of the Shamrock consortium. The consortium is important for a number of reasons: (1) it is a vendor and a user consortium, (2) it addresses a crucial user need, and (3) it illustrates a new level of maturity reached by the (still relatively young) document management market.
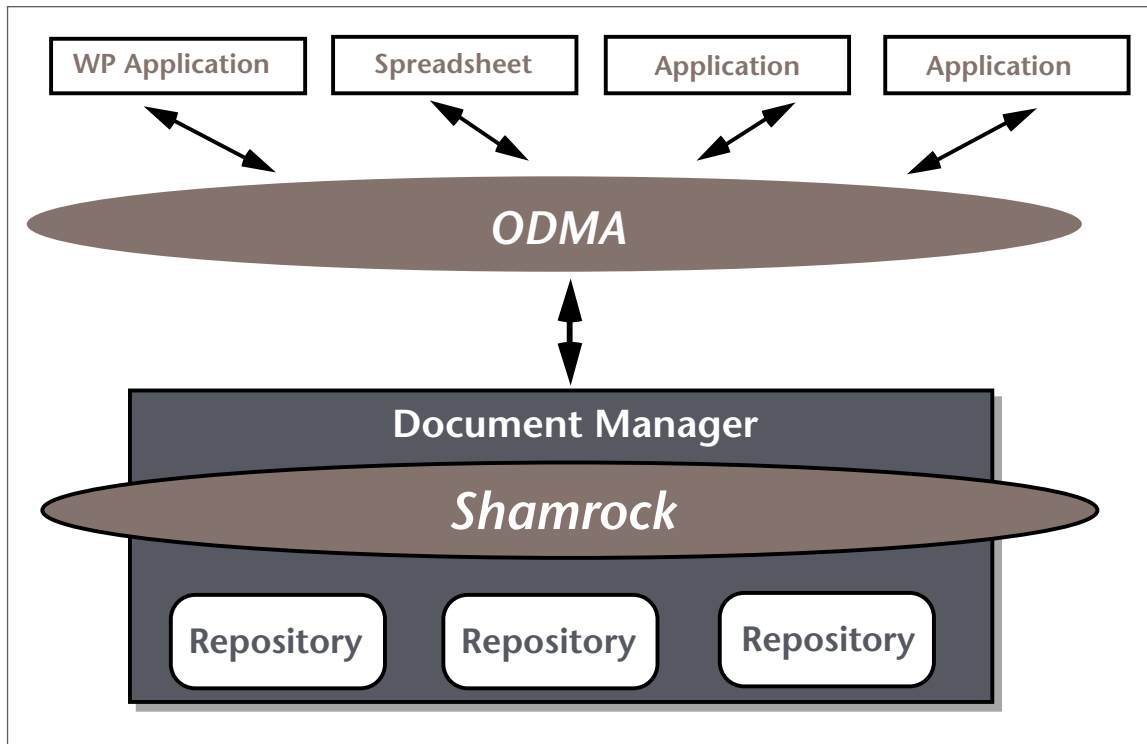
Shamrock was started by Saros and IBM, supposedly at a meeting on St. Patrick's day 1993, (thus the name). Some other members of Shamrock are: Interleaf, Microsoft, Coca-Cola, Aetna Life & Casualty, and Merck.

The basic premise behind Shamrock is that large enterprises are bound to have multiple document management applications just as they have multiple database applications today. As companies build multiple document repositories using different platforms and document management applications, they will sometimes need to be able to access documents in repositories created using document management applications to which they lack access.

Shamrock vs. ODMA

Shamrock is not the only consortium addressing this issue. ODMA (Open Document Management API) represents another attempt at a standard that ensures interoperability between document management systems and other software applications. Initiated by SoftSolutions, it also counts Oracle, Novell, Lotus, Documentum, and Watermark among it's members.

At first glance it looks as though Shamrock and ODMA are competing solutions to the same problem. This perception is further encouraged by the fact that Saros and SoftSolutions are competitors. Theirs, however, are not the only competitive issues at play here; there are many competing interests involved when formulating standards to address such fundamental business problems.

Figure 3

How

Shamrock &

ODMA Relate

(Sort Of).

There is, however, a difference in emphasis between the two approaches, as Figure 3 (taken from an ODMA document) attempts to illustrate. ODMA's objective is to provide a common API for common desktop applications (*e.g.*, spreadsheets and word processors) to get at information being managed by document management applications. Thus ODMA focuses on providing support for desktop "office" applications, while Shamrock focuses more on issues of repository management. Nonetheless, there is certainly quite a bit of overlap. Both organizations are attempting to play down their differences and there is talk of either: merging the two activities, or assigning different API calls to each group to keep the two activities from getting in each other's way.

Our view is that they should be merged, for several reasons. We suspect the overlap is pretty large; We don't think users of these proposed standards will want to have to track both initiatives; And we do not think a clean separation could last over time — the two groups would have to constantly work together to ensure they were not working at cross-purposes.

"Document Enabled NetWare"

To add to the confusion, XSoft and Novell made a joint announcement at Interop of "Document Enabled NetWare", a set of middleware tools to facilitate the sharing of documents and document management services across LANs. While it is easy to see that this meets some different needs than the Shamrock or ODMA APIs (*e.g.*, network printing services), again there is bound to be some overlap. (These issues may warrant a full analysis in a future issue, and will be among those highlighted at Documation '95)

Interleaf's Intellecte

Another announcement was important because it illustrates what is bound to be part of the overall trend for document management to penetrate more and more mainstream corporate applications. Interleaf's Intellecte is a packaging of products and services designed for a complete document management implementation in a fixed time frame at a well-defined cost. This kind of packaged solution approach, especially for "high end" document management, will make it less risky for managers to invest in what can be a very tricky integration exercise.

SoftQuad's Explorer and Hal's Olias

SoftQuad entered the electronic browser market with its announcement of *Explorer,* an SGML-based browser developed by Sweden's Synex Information AB. The browser is based on a database specially designed for accessing and viewing SGML information.

Meanwhile, Hal Software Systems announced an SGML based browser that is designed to provide a seamless interface to both in-house collections of SGML documents and information as well as World Wide Web (WWW) documents over the Internet. What makes these announcements interesting is that they demonstrate that the market for SGML-based viewers is perceived by vendors to be very strong — in fact, the number of new SGML-based browsers is keeping up with the new non-SGML-based viewers. And while Explorer is an example of the trend to integrate relational with SGML database technology, Hal's *Olias* teases us with a glimpse into a future where corporate information can be combined with the fast moving data on the "infobahn" to create integrated applications.

Folio's Conference Proceedings

Folio generously agreed to produce the Documation proceedings and show guide and to provide it at registration to all attendees on either a Macintosh or Windows disk. What makes this noteworthy is that it shows that it is not unreasonable to capture, package,

"We don't think users of these proposed standards will want to have to track both initiatives …"

and deliver electronic information in a very demanding time-constrained environment. All conferences are faced with the challenges of collecting materials from speakers and then building conference proceedings at the last minute that are useful. To date most still rely on thick collections of paper proceedings, or provide electronic proceedings months after the event. Folio was able to accept material up until just a few weeks before the conference and deliver the disks in advance. Of course, one still has the problem of presentation material not delivered at all, or of presentations created on a laptop the night before. But it is encouraging that we're moving in the right direction.

OTHER INDUSTRY EVENTS

It seems almost every conference covering some aspect of IT has some activity related to document management these days. This in itself may be the strongest evidence of the mainstream need for this document management functionality. We captured those that are of interest to our readers in our regular News section. Here, we focus on a few of the more significant announcements made at other shows this Spring.

Seybold

The annual Seybold Seminars in Boston, well attended as usual, featured a couple of announcements important for corporate document management applications.

WordPerfect's Envoy

WordPerfect announced 'Envoy' a portable document viewer directly aimed at Adobe's Acrobat, Interleaf's WorldView, No Hands' Common Ground and others. In overall functionality, it appears to resemble Acrobat more than the others, and the literature suggests it will compete on performance and resource requirements. We have not looked at it, so have no opinion yet about the product itself. But we think this was an important and revealing announcement. It is important because WordPerfect's size and resources mean the product is likely to have a strong impact on the market. It is revealing because it illustrates the growing importance of this kind of product, and because it suggests that this market will continue to attract new and larger players (as we argued in our analysis of the electronic viewer market in *Vol. 1, Num. 4*).

EBT's DynaBase

Electronic Book Technologies (EBT) used the Seybold event to announce their new document management product. DynaBase, is an SGML-based document database, that provides for the management of individual SGML elements, as well as SGML documents. Built on top of Object Design's object-oriented database, it provides upstream document management capability for EBT's DynaText viewer.

DynaBases's significance lies in its object-oriented architecture. While others have been working on SGML-based document management repository products using relational or hybrid (relational/object) database technology, EBT is the first to announce a commercial package based on a "pure" OODB. This is a direction that many have been talking about for some time.

SoftQuad and Quark

Quark and SoftQuad have announced an agreement where SoftQuad will develop a Quark XTension, SGML Enabler, to provide SGML support within Quark applications. Quark has historically been the most vocal opponent of SGML among suppliers of publishing systems and software. While not everyone in that community has fully embraced

SGML, virtually all of them have acknowledged it's importance. The announcement is probably the result of severe customer pressure on Quark. It is certainly good news for those Quark customers, especially large publishers, whose SGML requirements presented them with a difficult choice.

ArborText's PowerPaste

ArborText announced a new conversion product at Seybold that provides an interactive environment for conversion from word processing and publishing formats to SGML. While PowerPaste is far from the first SGML conversion tool, it is important in that it recognizes that human intervention is typically required for accurate SGML conversions, and emphasizes a non-technical interactive approach. PowerPaste also aids in the process of transforming SGML documents from one DTD (Document Type Definition) to another.

AIIM

The AIIM show (sponsored by the Association for Image and Information Management) is a huge trade show that attracted over 40,000 people this year. Although AIIM includes "document management" in it's description, it's main focus is on hardware for image storage and retrieval, and increasingly, as imaging vendors have added it, workflow software.

Xerox "Open Document Services"

Xerox chose AIIM to launch a mammoth presentation of their strategic direction. "The Document Company" wants to help businesses manage the entire document process. This means not only providing tools for copying and printing, but products and services for creating, managing, and distributing documents. "Distribute and print" rather than "print and distribute" is one of the (shortcut) ways they describe their philosophy. The idea is that information should be managed and distributed electronically, and printed locally, or wherever it would be most efficient.

Xerox wants to be a focal point for document distribution technology and services as we move toward a high bandwidth global communication infrastructure. They recognize they can't do this alone and have started forming partnerships with a huge number of companies to reach their goal. (There are dozens, including AT&T, Novell, EDS, and Interleaf.)

Xerox also announced DocuSP (Document Services Platform). DocuSP is basically an unbundling and opening-up of their DocuTech/DocuPrint technology. Both the hardware and software with APIs will be available for their Open Document Services partners to enable them to build integrated solutions.

Recognition & Partners

A telling example of how the document management industry has evolved recently is provided by the many partnership announcements made by Recognition International in the past few months. Companies included in the partnership blitz include: Oracle, Kodak, Documentum, and Saros! Here we have an imaging/workflow supplier teaming up with suppliers of (1) databases, (2) imaging/OCR subsystems, (3) enterprise compound document management systems, and (4) PC/file-based document management systems.

CONCLUSIONS

What became clear during Documation and at some of the other shows was that:

- The strength of the interest in document management of all types was confirmed. While we don't have any hard numbers to back it up, we think that interest in compound document management systems is picking up speed faster than interest in

image-based systems, although the distinction between the two is increasingly becoming less clear as (some) imaging suppliers add support for content types beyond images.

- Lots of confusion remains around document management and related technologies. But the confusion is now widely recognized and a dialog has been started among vendors, and that should help.

- We were struck by how much document management is being viewed as a basic requirement of enterprise-wide information and office systems. There is a recognized need to manage not only strategic documents (e.g., engineering drawings or clinical studies), but also office documents that play a supporting role in strategic applications (e.g., design review documents, correspondence records).

Just as electronic publishing features used to be found only in niche applications but are now expected in most office applications (e.g., multiple fonts in spreadsheets), now "image-enabled" PC applications are available, and soon, more advanced document management features (e.g., check-in/check-out control) will be everywhere, and expected.

- There will continue to be important niche applications for specific disciplines, (e.g., engineering drawings, and product information), but increasingly, managers will look at general purpose (and even "off-the-shelf" if available) solutions first to see if they will solve the problem.

- "Interoperability" is probably the one term that captures what the industry needs more than anything else. It's a big word and covers communications and networking issues as well as application integration and even information sharing. The Shamrock and ODMA initiatives are evidence of both the importance of interoperability and of mainstream interest in document management. The SGML Open consortium, whose members were out in force at Documation and at some other events, similarly generated a lot of interest.

- "Electronic delivery" in it's various guises continues as a driving force for more enhanced ways to manage documents and document information. "Documents on demand" or "just in time" documents whether in paper or electronic form require a system to manage the document information independently of the media upon which it is delivered on.

- Sufficient technology exists to build sophisticated document management systems to meet most existing business requirements. However, document management is a fundamental IT requirement, and in fact, is needed to supplement the inadequacy of file systems. We need, and will get, new file systems in the next couple of years.

Our View Of The Landscape

We are constantly helping companies (both vendors and corporate consumers) navigate around the document management landscape, and are often asked for a clear, concise definition of what document management is — we would love to oblige! A single definition however is unrealistic given the diversity of the market. On the other hand, it is possible to describe the landscape. Figure 1 describes the supplier market. Figure 4 is an illustration we sometimes use to get people oriented around product categories
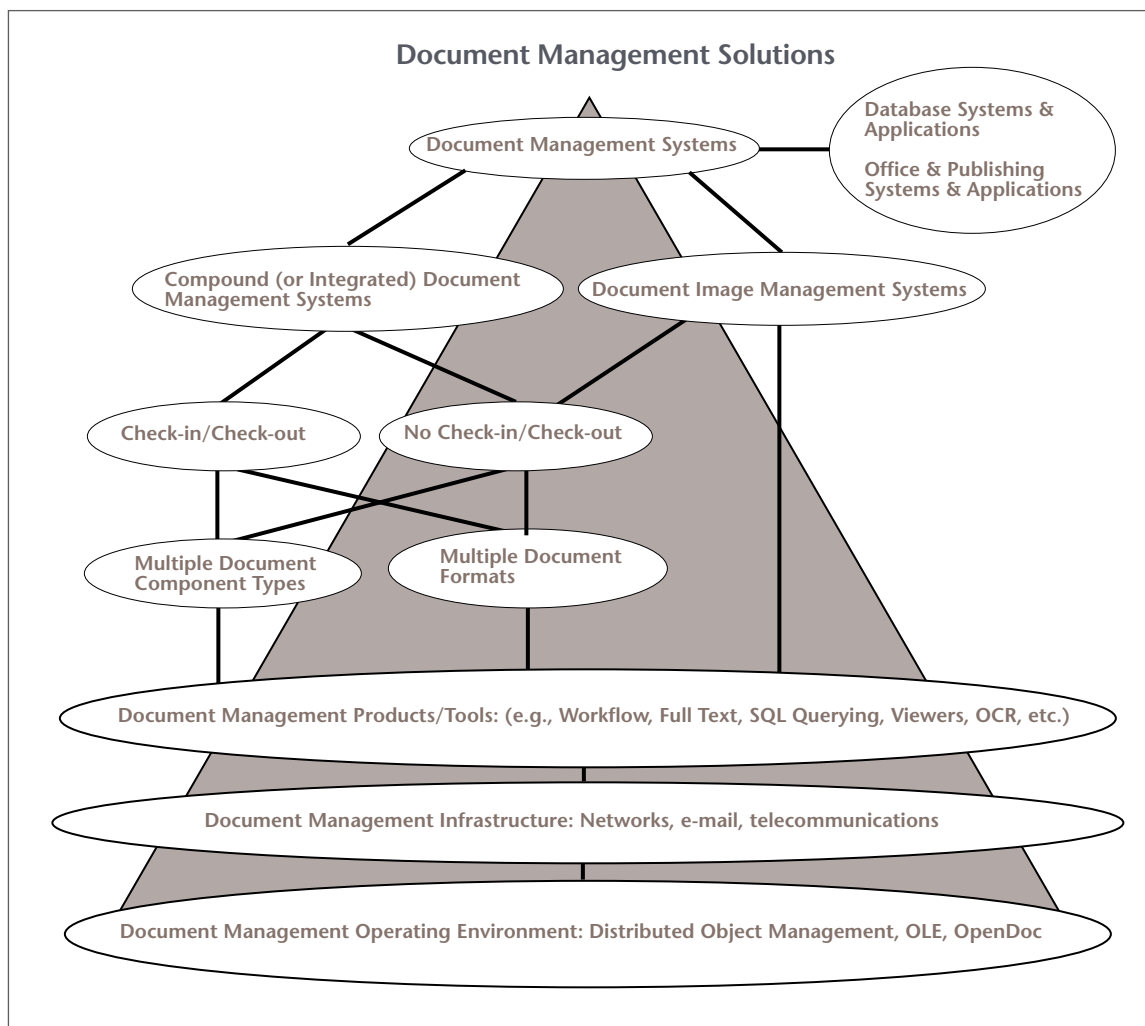
Figure 4

Product

Category

Landscape

Document Management Solutions

Document Management Systems

Database Systems & Applications

Office & Publishing Systems & Applications

Compound (or Integrated) Document Management Systems

Document Image Management Systems

Check-in/Check-out

No Check-in/Check-out

Multiple Document Component Types

Multiple Document Formats

Document Management Products/Tools: (e.g., Workflow, Full Text, SQL Querying, Viewers, OCR, etc.)

Document Management Infrastructure: Networks, e-mail, telecommunications

Document Management Operating Environment: Distributed Object Management, OLE, OpenDoc

SHANAHAN: "THEY GET IT!"

To me, the single biggest milestone of the Documation conference seemed to be the acknowledgment by end-users and information technology personnel of the need for document management. I found this to be an exciting phenomenon, especially since users have been trying for years to understand what "object-oriented" means.

Now that we're over that hurdle, the industry can move beyond trying to justify the value of document management and begin tackling issues of how to increase value to the enterprise. We can also begin defining standards and practices, rather than just dwelling on what might be.

"… the debate of relational vs. object-oriented repositories finally came to an end"

Another highlight: the debate of relational vs. object-oriented repositories finally came to an end. While it is possible to shave a square peg to fit in around hole, using a relational database as a document repository falls well short of a complete solution. Still, over the last several years, debate has raged over which technology, relational or object, is most appropriate approach for managing documents in a repository.

At Documation, the object paradigm finally won. Now the challenge lies in validating vendor claims to be object-oriented, because the majority of the products on the market today—even many that claim to be object-oriented—are, at their core, based on a relational model.

The other good news is that document management is finally reaching beyond the publications department. In the past, document management and applications have tailored their functions and features to meet the needs of the technical publications department. Not anymore. Delivering a document solution only for Tech Pubs is too little too late.

In many firms, documents either are the product or are integral to delivering the product. The information contained in these documents originates in engineering, marketing, research and development. To improve the quality and timeliness of documents, the information content must be accurately captured at the source and released in a controlled fashion.

As corporations today downsize or reengineer their processes to make them more efficient, many are flattening the organization and increasing the speed with which products are delivered to the marketplace. In the process many are capturing information sooner, opening the way for document management to be applied not just in Publications but throughout the corporation.

LAPLANTE: THE "INFO GAP" CLOSES

The longer information remains in digital form, the narrower the gap between that information and the people who need it. This enhances the timeliness and accessibility of documents, which in turn increases their value.

In the distribute-then-present model, presentation (formatting) is applied after data is distributed as a document, rather than before. As a result, data remains "live" for a longer period of time, resulting in higher-quality documents that are complete and accurate.

Digital distribution in this manner makes information much more accessible as well as more timely. Information becomes global rather than local — it can be poured into a document container when and where it is needed. And by bringing users closer to our information, we can provide them more opportunity to access it themselves, which drives up its value.

The new model presents challenges as well as opportunities. The need to maintain and update information becomes increasingly important, as does the requirement to ensure consistency across information and document databases. Security and data integrity are also issues when users are closer to information sources.

SGML, the international standard for document interchange, facilitates publishing from distributed data, since it provides more control over data integrity, ensures consistency of information across documents, and reduces the time required to develop new information products. A growing number of electronic document delivery systems can accept native SGML files directly, further streamlining the document distribution process.

One of the most significant benefits of SGML is the flexibility it provides. Since SGML separates content from format, SGML source files can be pumped out through different document preparation systems that support widely varying presentation media and platforms. SGML documents thus adapt readily to the display or print characteristics and requirements of different output devices — desktop, laptop, palmtop; paper, in U.S. and international sizes; and CD-ROM, network server, and Internet delivery.

What I found at Documation was broader acceptance of standards such as SGML, coupled with a willingness to reengineer processes to take better advantage of data in document form. Everyone looks for opportunities to reduce quantifiable printing and distribution costs, but more and more you hear firms citing ways of generating additional rev-

"… more and more you hear firms citing ways of generating additional revenues, rather than cutting costs, as more compelling reasons for re-engineering their document distribution systems"

enues, rather than cutting costs, as more compelling reasons for re-engineering their document distribution systems. One customer service organization developed two new information products as a result of bringing technical support documents on-line. Another speaker described how the transition to a system of digital distribution to remote printing sites enabled his company to increase the number of reports that were available for delivery and sale to its customers.

DOCUMATION '95 CALL FOR PAPERS

The deadline for submission of proposals for next years conference is July 15. If you would like to speak or have suggestions for other speakers please let us know as soon as possible. We are interested in corporate users with experience in implementing a document management system, especially those who have had interesting experiences integrating document management or electronic delivery with other corporate information systems (customer service, manufacturing, technical publications, *etc.*).

Proposals should be sent to: ptm@world.std.com or document@well.sf.ca.us or they may be faxed to (617) 576-5708 or (703) 548-2867.

If You Are An Exhibitor

There is still a fair amount of booth space available, but reservations are brisk and the best locations are going fast. To reserve your spot or for more information call Kaiser & Associates at (805) 984-4364, fax (805) 984-1870.

Industry News

In this section we 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 market-place. If you have news, keep us posted, and if you have comments on what you would like to see us cover, let us know.

SGML Open members vote for a Technical Resolution on Entity Management

The SGML Open sponsor members are currently voting on the consortium's first Technical Resolution. The final draft of Technical Resolution 9401:1994 on Entity Management issues has recently been distributed among SGML Open sponsors for a two month review and ballot period. If two-thirds of the votes accept the Draft Resolution, it will be approved as an official SGML Open Technical Resolution.

This Resolution addresses two different but related barriers to interoperability related to entity management:

A. that of resolving entity declarations for multiple vendor's applications on a given system, and

B. that of interchanging a set of files in such a way that appropriate entity identifiers can be attached to the proper files.

While it is well understood that there are many complex issues involved, this Resolution recommends a simple set of conventions that would address a useful subset of the complete problem. To address issue A, it suggests a format for an entity catalog that would handle the simple cases of mapping an external entity's identifier to a file name. To address issue B, it suggests a simple interchange packaging scheme that includes an interchange catalog that associates a public identifier to each file in the interchange package.

Once participating vendors implement the Resolution's recommendations, the SGML community should see a noticeable improvement in interoperability among various products in the area of external entity referencing and document interchange. It is expected that this Technical Resolution will be the first of an important set of consortium resolutions that will improve interoperability in the SGML market and extend the practical useability of SGML for the general use community.

Excalibur Technology To Be Used For Fingerprint Identification

What does this have to do with document management you ask? The technology being used is based on the same technology that Excalibur uses for text retrieval. Excalibur's EFS system is based on what the y call "digital pattern recognition" that can be used to recognize any type of digital information, including text, graphics, sound, video, etc. The fingerprint ID system will be one of the first (and most demanding) applications beyond text retrieval.

Separately Excalibur announced a new version of EFS at the AIIM show. Version 3.5 includes among it's new features a new client API.

SoftQuad ships Commercial Editor for the Internet

SoftQuad has announced SoftQuad HoTMetaL and SoftQuad HoTMetaL PRO graphical editors for the Internet. SoftQuad HoTMetaL integrates with Mosaic-the product of the US National Center for Supercomputing Applications (NCSA)-which makes navigating the Net easier by presenting a user-friendly way to view the text, graphics, movies and sound offered by the servers.

SoftQuad HoTMetaL is an HTML (Hypertext Markup Language) editor that supports native SGML. Since HTML is a subset of SGML, SoftQuad HoTMetaL automates the markup process and automatically validates all documents when they are loaded or saved.

SoftQuad HoTMetaL is a free, unsupported edition and will be available on June 1, 1994 from many of the same Internet sites and electronic bulletin boards as Mosaic (including NCSA). The first release will be for Sun Motif with a Microsoft Windows version following soon afterwards. SoftQuad HoTMetaL PRO is a fully supported commercial version with additional functionality that will allow users to handle more complex document structures. For example, it provides a table editor and offers more comprehensive support for authoring, such as user-defined macros and long or short sets of menus. It will be available before the end of June 1994 on Sun Motif and Microsoft Windows from SoftQuad and its Resellers. SoftQuad will also offer the HoTMetaL PRO product for sale over the Internet. HoTMetaL also includes an upgrade path to SoftQuad Author/Editor, which is a full cross-industry authoring environment that supports SGML.

Avalanche & SoftQuad Team Up Ease Access To The World Wide Web

Avalanche Development Company and SoftQuad Inc. announced a combined product suite that gives organizations a way to deliver documents using the Internet's information highway. The joint offering combines Avalanche's conversion technology with the new HTML editor, SoftQuad HoTMetaL. Avalanche's HTML product will be configured to serve as a HoTMetaL add-on that allows easy conversion from popular desktop word processing environments, including WordPerfect and Microsoft Word, into HTML. HTML (HyperText Markup Language) is the format that allows instantaneous display of documents stored on any computer attached to the World Wide Web.

Avalanche's product will be preconfigured to create HTML without additional configuration or setup so that users can begin publishing immediately. Avalanche will also offer a developer's version for modifying the application.

Fulcrum & WAIS Inc. Announce Strategic Partnership

This alliance will result in the integration of Fulcrum's SearchServer with the WAIS Network Publishing Protocol. Fulcrum plans to release the product at the end of 1994. There was no mention of exclusivity in the agreement. Presumably WAIS will also continue to provide it's own retrieval technology.

PC DOCS Licenses Imaging Software From Watermark

PC Docs announced that it has licensed Watermark technology for inclusion in it's PC DOCS OPEN document management system. Until now PC DOCS focused on Windows and DOS word processing files.

Saros Announces Two Alliances With Workflow and Imaging Vendors

Saros has aligned itself with Watermark and Recognition in agreements that call for them to provide integrated solutions to their customers. No specific products were announced. Notice that they have picked both a "high-end' production supplier and a "low-end" desktop vendor.

Lotus Beefs Up SmarText for Electronic Publishing

Lotus announced the release of SmarText version 3.0 at the Seybold conference in Boston. The new features, and positioning, is specifically aimed at the corporate electronic distribution market. Included in the new release is a way to link SmarText documents to Notes documents.

EBT Enters Conversion Software Market.

EBT, in addition to announcing a new OODB-based document management system (covered in our main article), also announced DynaTag. DynaTag is designed to let non-technical types convert WP and publishing system files and Rainbow SGML documents into SGML-based DynaText electronic books. The interactive approach is similar to ArborText's PowerPaste.

ArborText Ships Windows SGML Software

ArborText SGML-based editing and publishing software for Windows is now shipping. This is significant since it adds an important choice to those who want SGML but don't want to invest in a workstation to get there

Interleaf Announces WorldView 2.0 Ships and New Partnership With Berger-Levreault

Interleaf announced a new version of WorldView and released (WorldView) sales figures for fiscal year 1994 of $14 million. There are a host of new features including SGML support. WorldView2 accepts input of Standard Generalized Markup Language (SGML) files and can output the entire original SGML text file, making the information reusable.

Under the agreement with Berger-Levreault/Advanced Information Systems, Interleaf will integrate the SGML/Store database system with its Relational Document Manager (RDM) product.

Interleaf said it will begin delivering the SGML/Store technology immediately to its customer base and will integrate the SGML repository into its next major release of RDM which is planned for 1995.

WordPerfect Builds OpenDoc Parts

Wordperfect has announced that it has built OpenDoc parts for the Macintosh versions of their Draw and Chart functions. WordPerfect said that they also have both OpenDoc parts running in OLE (Microsoft's Object Linking and Embedding technology) on a Windows machine. This underscores WordPerfect's stated intention to actively support OpenDoc.

ORACLE STARTS TALKING ABOUT 'ORACLE DOCUMENTS'

While it's no secret that Oracle is keenly interested in the document management market, and in the 'groupware' market (*i.e.*, Lotus Notes), they are stepping-up the volume and are referring to a new product to be called 'Oracle Documents'. What is unclear at the is point is how much document management, as opposed to 'groupware', features will be included.

WANG RELEASES OPEN/PROFOUND

Wang announced it has started shipping a new version of it's document management software. Open/profound support file types other than images and runs as a client on Windows machines. The new pricing makes it competitive with other PC-based document management products like PC-DOCS.

XYVISION UPDATES PDM

Xyvision has released version 2.1 of its Parlance Document Manager (PDM). The major additions are a PC Windows client, and the integration of INFORMIX-Online as the underlying database.

IDI TEAMS UP WITH CAP GEMINI

CAP Gemini Sogeti, by signing-up with IDI joins other large integrators in the document management market. The agreement calls for CAP Gemini to provide implementation services and support using IDI's BASIS plus family of document management and retrieval products.

WORKGROUP ANNOUNCES NEW VERSION OF CMS/PRO

Workgroup Technology is coming out with an enhanced version of CMS/Pro (version 2.0) that provides additional version control, release cycle management and configuration management. Release dates were not provided.

FILENET TO PORT TO SUN

FileNet has signed a letter of intent to port its document imaging and workflow software to Sun's Solaris/Sparc platform. The new software will start to be available in the first quarter next year.

FileNet separately announced a new version of their technical documentation (engineering drawing) viewing and markup software Revise 2.0 for Windows. The new version will be available in June.

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.

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

Workflow. August 11-12, 1994, San Jose CA. Call (800) 2470262, Fax (602) 661-0449.

CALS Pacific. August 30 - September 2, Taipei, Taiwan. Call (703) 578-0301 or +49 30 882 6656, Fax (703) 578-3386 or +49 30 883 8811.

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.

Seybold Computer Publishing. September 13-16, 1994, San Francisco, CA. The enormous annual computer publishing exposition and conference. Call (800) 433-5200.

CALS Japan '94. September 26-29, 1994, Tokyo, Japan. The first CALS conference in Japan. Will focus on commercial CALS applications. Call (703) 578-0301 or +49 30 882 6656, Fax (703) 578-3386 or +49 30 883 8811.

EDMS Vision '94. September 26-29, 1994, Anaheim, CA. Conference and show devoted to engineering document management systems. Call (800) 242-6822.

SGML Singapore '94. October 10-12, Singapore. The first SGML event sponsored by the GCA in the Pacific Rim. Call (703) 519-8160, Fax (703) 548-2867.

PTM Document Management & Electronic Delivery Seminars. November 28-29, London, UK. These two day seminars are conducted by PTM & Gilbane Report staff and are managed by Technology Appraisals. Call +44 81 893 3986 or (617) 576-5700, Fax +44 81 744 1149 or (617) 576-5708.

Xplor '94. November 7-11, 1994, Phoenix, AZ. The large 15th annual global gathering of electronic printer users and vendors. Call (800) 669-7567, Fax (310) 375 4240.

SGML '94. November 7-11, Vienna, VA. *The* annual SGML event in the U.S. for both novice and advanced SGML users or developers. Call (703) 519-8160, Fax (703) 548-2867.

CALS Expo '94. December 5-8, Long Beach, CA. The annual expo and conference covering CALS activity in the U.S. and internationally. Heavy defense industry emphasis. Call (202) 775-1440, Fax (202) 775-1309.

PTM Document Management & Electronic Delivery Seminars. January 18-20, Location TBD (either Geneva, Switzerland, or Amsterdam, The Netherlands). These two day seminars are conducted by PTM & Gilbane Report staff and are managed by Technology Appraisals. Call +44 81 893 3986 or (617) 576-5700, Fax +44 81 744 1149 or (617) 576-5708.

Documation '95. March 7-9, Long Beach, CA. The conference and exposition of the year for the document management and document computing industry. Covering all aspects of enterprise document management applications. Co-sponsored by PTM, The Gilbane Report, the GCA, and the GCA Research Institute. Call (703) 519-8160 or (617) 576-5700, Fax (703) 548-2867, or (617) 576-5708.

LETTERS FROM OUR READERS

We encourage letters from our readers on any topic related to the areas we cover in the report, and are especially interested in opinions about articles we publish. So let us know what you think! If you don't have time to write a formal letter, send us a quick note or fax (email preferred).

SGML OPEN COMMENTS ON ANDERSEN'S VIEW OF SGML FOR TEXT 2000

6 May 1994

Ref: TEXT 2000 - Gilbane Report, March/April 1994

We read this article with interest, and a high proportion of confusion, especially the sections entitled "Longer-Term Benefits" and "Lessons Learned, SGML". We are concerned that this article, much like the system it describes, focused almost exclusively on technology and gave little consideration to the information being processed. It adopted a proprietary solution for the encoding of the information. This is unusual and not good design practice when the value of the information far exceeds the value of the system.

Our confusion arises from the statements in "Lessons Learned, SGML" that the "...system had to handle highly structured documents with complex tagging requirements...", that "...users have the ability to tag text for use in document shredding, searching and composition.", and that "...adopting it [SGML] instead, would require more knowledge about document structure, etc., than a typical TEXT 2000 user has." These statements appear contradictory in that being able to know what and how to "...tag text for use in shredding, searching and composition" can't be far different from having "...knowledge about document structure..." The implication of the article is that tagging text is fine, but that using SGML is not, when in reality the use of a good SGML tool that is properly configured can constitute a structurally-aware tagging application that would require even less knowledge of the document structure and format than a non-SGML application.

This same portion of the article mentions that although the tagging requirements would typically suggest an SGML solution, "...this requirement had to be balanced against a requirement that the drafting tools be easy to use and intuitive to the end-user." We would like to point out that ease of use and the adoption of SGML are not contradictory. As is true with applications in any area of technology, one can create less-than-intuitive SGML applications. However, that is the fault of the tool and/or application designer, not SGML. In fact, SGML's decoupling of the data format that encodes the structure of the information from the presentation and end-use of the information (to say nothing of the advantages of its system- and vendor-independence)encourages the development of easy to use and intuitive interfaces. What is required is that tools that are powerful and flexible enough to support SGML properly are used as part of a system that is carefully designed and configured in a way that is optimized for the task and user community at hand. Given everything we read in this article, we are surprised that the project team decided that they were not capable of providing an intuitive and accessible system using SGML. We feel that there are many resources, including those available from our member organizations, who could have helped the project team to do so.

Lastly, we were disappointed to see the list of Longer-Term Benefits. They focus solely on the system and not on the information being used within it. Of note here is the argument for future compatibility of system components. The article states that it "...allows for the changing or updating of individual pieces of the system over time rather than replacing the entire system." This may be true to a certain extent. However, if the new component can't use the proprietary encoding used within the system, the addition of the new component will mean a costly data conversion process. Worse yet, the new, desirable component may have to be foregone because it is not feasible to re-encode the data. When the next generation of more-user-friendly tools arrives, it will mean a massive data conversion task to take advantage of the technology. Reconverting data at today's dollar value is bound to be cheaper than doing it at tomorrow's dollar value. Compound that with the increase in volume of data to be converted and TEXT 2000 may in the longer term be judged a technological wonder, but a missed opportunity.

Board of Directors
SGML Open
910 Beaver Grade Road, #3008
Coraopolis, PA 15108

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?

Vol. 2, No. 1.

State Of Wisconsin Legislature TEXT 2000 — Reengineering For Document Management.

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 PTM Document Management & 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.
University Place, 124 Mt. Auburn St., Suite 200N , Cambridge, MA 02138
Fax: (617) 576-5708 • To order by phone call: (617) 576-5700

HOW TO FIND OUT MORE ABOUT COMPANIES MENTIONED IN THIS ISSUE

There are too many companies mentioned in this issue to provide addresses for all of them. If you would like contact information give us a call or send us a fax or email note (definitely preferred!) at our new address listed below and we'll provide the informa-

OUR NEW ADDRESS

PTM and The Gilbane Report have moved into new and larger offices in Harvard Square. Our new address is:

Publishing Technology Management, Inc.
University Place
124 Mt. Auburn Street, Suite 200N
Cambridge, MA 02138
Tel. 617.576.5700
Fax 617.576.5708
Internet: ptm@world.std.com