

THE GILBANE REPORT™

Content, Computing, and Commerce – Technology & Trends

Vol. 10, No. 8
October, 2002

www.gilbane.com

Published by:
Bluebill Advisors, Inc.
763 Massachusetts Ave.
Cambridge, MA 02139
USA
(617) 497.9443
Fax (617) 497.5256
www.bluebilladvisors.com

Editor:
Frank Gilbane
frank@gilbane.com
(617) 497.9443

Editors Emeriti:
Tim Bray
tbray@textuality.com
(604) 708.9592
David Weinberger
self@evident.com
(617) 738.8323

Associate Editors:
Bill Trippe
bill@gilbane.com
(617) 497.9443
David R. Guenette
david@gilbane.com
(617) 868.6093

Recent Contributors:
Sebastian Holst
sebastian@gilbane.com
Girish Altekar
girish@deepcoolclear.com
Bob Boiko
bob@metatorial.com
JoAnn Hackos
Joann.hackos@comtech-serv.com
Rita Warren
rita@ziacontent.com

Associate Analyst:
Jared Spataro
jared@gilbane.com

Contributing Analyst:
Bryan Crow
bryan@gilbane.com

Production Assistant:
Sarah G. Dionne
sarah@gilbane.com
(617) 497.9443

Subscriptions:
sales@gilbane.com
(617) 497.9443

Customer Service:
customerservice@gilbane.com

Consulting Inquiries:
consulting@gilbane.com
(617) 497.9443

Gilbane Report White Papers:
Jared Spataro
jared@gilbane.com
(617) 497.9443

Speaking Engagements:
speakersbureau@gilbane.com
(617) 497.9443

THE ROLE OF XML IN CONTENT MANAGEMENT

The management of semi-structured or unstructured data has always depended on markup languages. Before the Web it was SGML or proprietary markup languages, now it is XML. This dependence was mutual – in practice, unstructured information management (mainly for publishing) was the *only* use of markup. However, the wild success of XML is due to its acceptance as a way to encode and share *all* kinds of structured and unstructured data, including code. Ironically, advocating XML for content or document management was actually disparaged by many early XML evangelists because they were afraid XML would be seen as being limited to publishing-oriented applications. This was in spite of the fact that, while most XML development was targeting application integration, most deployment was for content applications.

Today, you wouldn't implement a content management solution without thinking very carefully about what role XML should play. Should it be used for content, for metadata, for application integration, for information integration? Where in the create/manage/deliver cycle should it be used? Where do Web Services fit in? What about WebDAV? Contributor Lauren Wood returns this month with a look at how businesses are actually using XML in content management implementations, and how they view XML's role in the future. Lauren's report will provide you with an outline to help you organize your thoughts about the role XML should play in your content management implementation.

CONTENTS

The Role of XML in Content Management.....	2
Industry News.....	9
Subscription Form & Calendar	20

THE ROLE OF XML IN CONTENT MANAGEMENT

EXECUTIVE SUMMARY

XML is an extremely flexible technology that can fulfill several roles in any software application. Content management is no exception to this. This survey discusses some of the roles that XML can play in a content management system (CMS) and whether there is much industry support or customer demand for such support.

On speaking with several people representing companies and customers in this area, I found that there is increasing demand and good support for XML for content; less support or demand for XML metadata support; increasing support but little current demand for Web Services, and some support but more demand for WebDAV.

INTRODUCTION

One of the interesting things about XML is that the principles behind it are so simple that it can be used in many different ways. If we ignore all the other specifications and concentrate for a moment on the simple XML 1.0 specification¹, we see that what XML does is give us a way of labeling information. This labeled information is relatively easy to process, and is readable by humans (depending on the choice of the labels). Most of the 30-page specification is taken up with defining the syntax to make these two important facets possible (along with allowing for graphics, internationalization, and robust error-handling). Since XML can be used for so many different things it's not surprising that it is used in many different roles in the content management world as well.

This article will talk about "content management" or "CMS" (content management systems) and include all the variations that are appropriate, such as information management, knowledge management, or document management. Yes, these are all different. In terms of where XML can be used, however, they are similar enough to justify lumping them all together under one label.

XML use in a CMS can be divided into two main categories:

- XML for content (including metadata)
- XML for plumbing (including Web Services)

Several people in the CMS business were interviewed for this article and I asked them about the current use of XML in both of these categories. The results were interesting and show that XML is starting to push past the hype into the mainstream. Opinions varied wildly as to how widely XML will be used in the near fu-

¹ Found at <http://www.w3.org/TR/REC-xml>

ture, and for what; the synthesis presented in this article is my own and should not be attributed to any of the people I spoke to.

XML FOR CONTENT

The origins of XML are well known: it is a streamlined version of SGML (Standard Generalized Markup Language). SGML was particularly well suited to being used for technical documentation and publishing. The concepts that led to SGML being used for hard documentation problems are still present in XML. For example, the airline industry developed methods for coping with the fact that every airplane is individual and needs a maintenance manual that includes all the work that has been carried out on that particular plane. Such methods require a sophisticated view of the documents incorporating relevant metadata (which airplane it is) and content (what needs to be done), as well as information as to workflow (due date or time the job needs to be done by, and which team does the maintenance) and integration to other systems (who gets billed for the work). XML is the only common content format that readily allows for such sophistication.

Obviously, such sophistication is not needed for every document or for every company. But even smaller companies with less extreme needs still want to be able to repurpose content for print, web, or other formats and many are turning to XML for this. This content needs to be managed and so the demand is rising from customers for a CMS that can handle XML well enough for their needs.

Usually customer requirements are a mixture of three basic needs: reusing content, repurposing content, and keeping their content independent of the applications used to create and manage it.

- Content is *reused* when one it appears in more than one context. A common example is a copyright statement that may appear in hundreds of separate documents. If the statement is updated, the change will immediately appear in each of the documents that contain the copyright.
- *Repurposing* content means delivering that content in more than one format or medium. The most common repurposing need is to deliver information in both HTML and print (often PDF).
- *Application independence* has a number of different implications. Most often, it means that an organization will not be locked into a particular vendor. In addition, different departments within the same enterprise can adopt an XML model even though they may have differing systems in place.

In the current economic climate, companies are being much more careful about where they put their money, and much more cognizant of the need for a technology strategy plan. This means they will probably implement systems that better suit their needs. There appears to be an upturn with companies doing feasibility studies and pilot projects, ready for implementing in the next 6-12 months. And many of these projects will be using XML for content. Not many of these projects are in the large enterprise content management space; we're seeing more departmental projects, or projects for particular types of documents. HP exemplifies the type of company implementing the latter – HP has many dif-

ferent product groups that all produce documents for technical support or product catalogs. It makes sense to use one strategy and one type of system for all of those documents, no matter which department produces them. True enterprise-wide content management is still some time off, though there are some Fortune 200 companies looking at centralizing their information flows to allow for enterprise-wide access.

XML for content is often thought of principally as a technology used in publishing. The “traditional” publishing industry that started with SGML is moving to XML because of the cheaper tools, and often incurring some expense in moving their content to obey the stricter XML syntax rules. In general, however, they understand what XML is good for and have for some years. What is interesting now is that many other industries are also moving to XML without having a background in SGML. For example, Web content management systems that use XML content and then transform on the server to HTML or PDF are increasingly popular with smaller companies from a multitude of industries seeking an easier way to maintain their web sites.

One of the biggest areas of growth for XML is e-learning. Demand for e-learning is growing fast, and from multiple directions. Students at colleges and universities are increasingly expecting material to be available online to supplement their lectures. Adults are upgrading their qualifications in online courses, or expect online support for those courses they take in evening school. And companies are running training for their employees and their customers online to avoid travel costs and disruptions.

Cisco uses XML for an e-learning system that they use for employees and for customers. There are two major reasons for using XML.

1. The engineer who knows how the new switch or router works only has to write it all down once. The content can then be used to create derivative works, such as for marketing materials, without having to go back to the engineer. Prior to using XML, the engineer was a bottleneck, because everything had to be authored by that person (which also meant s/he couldn't do anything else!)
2. The content can be tailored to the needs of the person receiving the training. Adults in a 4-day course with 10 years of experience have different needs to college students who have 4 months to learn the same material, but have no experience in the area.

The companies using XML together with a CMS range across the spectrum, from Fortune 200 to small. Companies are in publishing, in finance, in manufacturing. Consumer products companies such as Kohler and Proctor and Gamble are implementing XML systems as part of their business processes, realizing that the documentation related to what they are selling to consumers now has to be delivered in a variety of formats, be accurate, and be timely. As companies increasingly sell into markets outside of their home territory they are also finding they need to produce different documents to go with those products. There may be different products in different countries, different names, or different marketing approaches – not to mention different languages! The recent announcement of significant XML support in Microsoft “Office 11”, along with the XML support already available in Corel WordPerfect and Sun StarOffice, shows that XML for content is reaching the mainstream.

In fact, the number of customers using XML for content now has reached the stage where any CMS vendor that hasn't already implemented XML support probably can't – presumably because of some problem in their underlying technology. For the customer shopping for a CMS, that means figuring out where XML will be used in the business process, and what other data formats will be stored as well. The days are past when a "silo" mentality of storing XML in an XML content store, and other documents in some other store, made sense. This is no longer necessary. All the larger CMS vendors support multiple data formats, including XML, so you can store your Word files, XML files, and multi-media in the same facility. Whether you should store everything in one repository, or whether you should have multiple repositories, depends on a number of factors that no longer need to have anything to do with the format the content is stored in. Over the last year or so we've seen a lot of movement in this space, between the more traditional CMS vendors adding XML support (sometimes you need to get an optional module to get all the features, such as chunking), the relational database vendors adding some degree of check-in and check-out, and the Web CMS vendors broadening their format support to include XML and office document formats.

With all this competition, the prices are coming down and the vendors and consultants are keen to get business. This makes it all the more important for companies thinking of installing a new CMS, or updating an existing one, to know what sort of content they wish to store. When they know that, then they can figure out how much XML support they need, and they can shop that around to the vendors and consultants. Knowing what you need is necessary – for example, one big variation between products is in how efficient they are at finding and checking out a chunk (portion) of an XML document when it's stored in the CMS. Some products are much slower than others at finding or checking out the chunk when there are many very small chunks; this should only worry those who need such high granularity for their XML.

Authoring Content

So what's the biggest problem with XML content? Authoring it... The authoring tools are becoming more capable and people are starting to figure out that the ease of processing XML content can outweigh the pain of creating it, but there is still some way to go. Since XML is so flexible, any XML authoring tool needs to be configured to match the schema and should also be configured to match the author's needs and knowledge. This, in a sense, is the "last mile" issue for the XML content industry. Frequently, the last issue considered in a well thought out XML system is the content creation process. However, a lot of good work and otherwise admirable effort can be undermined if the ease of use of the system isn't carefully considered. Small changes to the data model and authoring tool user interface or configuration can often produce dramatic improvements in productivity and quality.

Metadata

Metadata is the connecting tissue for all CMSs. It tells the CMS what the content is, who created it, who may read it, who may change it, where it fits in the workflow, and what sorts of operations may be performed on it. Metadata can do more, however. If the Semantic Web ever becomes reality (even if it never quite reaches the grandiose dreams some people have) it will be because sufficient metadata has been added to each bit of relevant content.

Metadata can be stored as XML, as indexes in a relational database, or in some CMS-specific storage format. For some purposes, the format it is stored in is irrelevant. Metadata that is more volatile than the underlying content, such as stage of a workflow process, or date the item moved from one stage to the other, is often stored outside of the XML. An XML format becomes useful in other scenarios, such as integrating different systems, or if the metadata is complicated enough to warrant storing it in a rich hierarchical format. In particular, a rich taxonomy provides a way to navigate through content following different navigation paths.

Since integration of different CMSs, passing around content complete with the metadata, and the requirements for rich, hierarchically structured metadata are just starting to become important for many people, the various metadata standards (in which I include topic maps and RDF) have not yet experienced the updraft that XML for content has. Metadata is the second layer of a complete content management system and requires at least as much thought as the design of a document schema for authoring in XML does. At this stage in the technology cycle, there isn't yet the experience in metadata system design that there is in document modeling; the best practices (which depend on the industry) are still being worked on. Metadata is hard: Mark Hale estimates that to fully classify a single document requires 60-90 minutes of human thought. Automatic metadata generation can help, but it will be some time before it's satisfactory.

Thus metadata is another area where the customer requirements document must be fully fledged out. Is the metadata required simply for workflow and basic search? Or will the content be passed around between divisions, or even between companies? If the latter, an XML format may be the right answer. If so, is there an applicable metadata standard or ontology that could be used?

XML FOR PLUMBING

Web Services

Web Services has a hype factor that rivals that of XML a couple of years ago. The number of articles proclaiming the virtues of XML, and the number of products proudly claiming XML prowess have decreased, simply because XML is now mainstream and all CMSs are *expected* to support it. The number of articles about the virtues and problems of Web Services has increased to fill that void.

Web Services² is an example of XML plumbing. The configuration files that determine how a piece of content is passed from one system to another are written in XML (actually a subset of XML). Web Services at the moment seem to be more hype than reality, but the economics of technology are such that there's a good chance that Web Services will become a basic part of systems infrastructure in the next two years or so. It will be used for passing around information between systems and thus for integration. Web Services are a relatively easy addition to most CMSs so there is push to implement from the vendors as well as the analysts who are writing all those articles mentioned in the paragraph above. The standards development isn't quite ready for primetime yet; some of the impor-

² We are talking about Web Services based on the W3C standards (SOAP *etc.*). Sometimes the term is used in a much broader way.

tant pieces such as security are still being worked on, but the basic shape is taking place. There are still some technical hurdles as well, such as the fact that SOAP only supports a subset of XML; various ways to solve this problem are also being worked on.

Do customers really want Web Services? Some do, depending on their corporate tolerance for risk or the technical vision of the person in the CTO office. I'm hearing far more about companies looking at adding Web Services support to their technical strategy over the next two years or so than wanting to add it immediately, though some who enjoy being on the bleeding edge are implementing it already. A large part of this planning is because companies need to integrate various systems. At the moment many (mostly the larger companies) are using J2EE for integration while many others (smaller to mid-size companies) are looking at migrating to .Net from COM. Web Services will be an important part of both of these platforms and so it makes sense to make sure components of an overall strategy, such as the CMS, also support the appropriate methods for integration. Web Services should enable integration between the J2EE and the .Net worlds; it remains to be seen just how robust and with what performance this integration can be carried out in the real world.

WebDAV (Web-based Distributed Authoring and Versioning)

Another piece of the puzzle that uses XML as plumbing, WebDAV is a relatively unknown specification that enables lightweight content management. It functions as a set of extensions to the web protocol HTTP (unlike Web Services, which can also function via other protocols such as email). These extensions are defined using XML. WebDAV (often called DAV for short) allows for basic CM functionality such as locking and metadata assignment; versioning is still being developed. It is not sufficient for a full-blown, all-the-bells-and-whistles CMS, but adequate for a lot of smaller uses where all the features of a large, expensive CMS are not needed. Once versioning has been added to WebDAV so that the basic check-in and check-out is supported, it will do much of what small groups of people need. There appears to be some customer demand for WebDAV in various tools such as XML authoring tool vendors, so that they can implement their own basic CMS. The larger CMS vendors are also implementing WebDAV (though the implementation isn't always supported) to enable a basic level of automatic integration with other tools without having to write special custom integrations for every tool on the market that a customer might want to use with the CMS. For many vendors, of course, there isn't the same level of urgency to implement WebDAV as they already have integrations with their favored third-party tools using their own methods. Customer demand seems to be having the desired effect, however.

MORE INFORMATION

Many of the topics discussed in this article will be discussed in much more depth at the forthcoming XML 2002 Conference and Exposition, to be held in Baltimore, Maryland in the week of December 8-13. Many of the people I spoke to in researching this article will be speaking at the conference on content management, metadata, and Web Services. There are also Town Hall meetings on these topics that give a forum for in-depth questions and discussions. The exhibit space includes many CMS vendors who will be showing their XML support. More information is at <http://www.xmlconference.org>.

(Note that full Gilbane Report subscribers **Save \$300 off the cost of a Conference Gold Pass**. Login to the Gilbane subscribers section at www.gilbane.com to get the discount priority code to use on the registration form. Discounts cannot be combined. – ed.)

ACKNOWLEDGEMENTS

I would like to thank everyone who spent time talking to me about the role XML plays in the content management world. I very much appreciate the input and the insights that they gave me. I spoke with Brian Buehling, Dakota Systems; Chris Wolff, Thomson; Jay di Silvestri³, Corel; Jay Todtenbier, Cisco; Jon Parsons and Rich Pasewark, XyEnterprise; Lubor Ptacek, Documentum; Mark Hale, Interwoven; Mike Champion, Software AG; Ron Daniel, Taxonomy Strategies; Sebastian Holst, Artesia; Todd Price, Stellent.

Lauren Wood
Lauren@textuality.com

³ Thanks also to Jay for proofreading.

INDUSTRY NEWS

More recent news, old news (to January 1999), and commentary is available at www.gilbane.com/

INXIGHT ANNOUNCES SMARTDISCOVERY 3.0

9/30/2002

Inxight Software, Inc., announced Inxight SmartDiscovery 3.0. Inxight SmartDiscovery expands upon Inxight's existing metatext extraction and analysis applications, adding features that include taxonomy management, enterprise-class categorization and a guided information retrieval environment. In addition, Inxight SmartDiscovery provides text analysis and retrieval capabilities that include on-the-fly entity extraction that identifies names of people, places, things and relationships from documents and groups them by category; automatic document summarization that creates intelligent summaries in a fraction of a second; a similar document finder that provides all other similar documents for a given document; concept search that organizes query results into a hierarchy of shared topics and themes; full text search that features basic keyword search and a Boolean information retrieval model for specifying the exact information needed; and multi-language support for accurately analyzing and retrieving information in 23 languages. Inxight SmartDiscovery is available immediately. www.inxight.com

ORACLE ANNOUNCES ALLIANCE WITH MOHOMINE

9/30/2002

Oracle Corp. announced an alliance with Mohomine, Inc. The integration offers mutual customers a streamlined process that reduces the number of resources needed to reformat and enter resume data during the recruiting cycle. Integrating with Oracle iRecruitment and the Oracle E-Business Suite, enables Mohomine to deliver resume content parsing for documents written in various styles and formats from within a single human resources environment. Mohomine's Resume Extractor automates and simplifies the process of attribute extraction from resumes by utilizing pattern recognition technology with "learning-by-example" techniques from within a single system environment. The Oracle and Mohomine solution is able to: accept various input formats including Word, PDF, Text, HTML, RTF and email/POP access; output resumes into parsed XML, using HR-XML standards; and, integrate with Oracle iRecruitment through the Internet. www.oracle.com

PROTEGE & VOQUETTE MERGER LEADS TO LAUNCH OF SEMAGIX

9/30/2002

Following a merger between Protege Ltd and Voquette Inc., a new enterprise software company called Semagix is being launched. The new company's semantic metadata management technology "Freedom" enables organizations to classify, manage and intelligently exploit structured and unstructured content from any source. This delivers information discovery capabilities and comprehensive real-time content analysis. Under the terms of the agreement U.K.-based Protege Ltd has merged with San Mateo-based Voquette Inc. in an all-stock deal. www.semagix.com

IXIASOFT & ADOBE JOIN FORCES

9/30/2002

IXIASOFT and Adobe Systems Incorporated announced that they will join forces to co-market and integrate IXIASOFT's TEXTML Server with Adobe FrameMaker 7.0. The initiative encompasses a broad base of marketing activities aimed towards the aerospace and defense industry, concentrated mainly on the synergy that exists between TEXTML Server and FrameMaker 7.0. Development efforts are moving ahead to provide seamless integration between the two products and will enable users to produce XML content in the environment provided by FrameMaker and store and publish this content using TEXTML Server, thus creating a complete document workflow system from production to searching and publishing. www.ixiasoft.com, www.adobe.com

eXCELON ANNOUNCES RELEASE OF eXTENSIBLE INFORMATION SERVER & STYLUS STUDIO

9/30/2002

eXcelon Corporation announced the availability of eXtensible Information Server (XIS) 3.12. XIS 3.12 facilitates XML-based interoperability across .NET and J2EE environments while delivering faster access to data in XML business documents with XQuery support and Verity full-text search. XIS is an XML database that allows XML business documents to be dynamically extended while providing granular access to elements of information contained in the documents. XIS now manages XML business documents in the .NET environment, to complement existing support for J2EE environments. XIS promotes heterogeneous interoperability between software platforms; companies can now use XML business documents as the foundation of composite business applications that integrate systems based on both J2EE and .NET architectures. XIS 3.12 will be available within 30 days. They also announced the release of Stylus Studio XSLT Integrated Development Environment (IDE) 4.5 to enable development teams to deliver applications based on XML and XSLT faster by enabling developers to generate XSLT stylesheets automatically from HTML pages, providing advanced debugging support including XSLT debugging for SAXON and .NET processors. Stylus Studio 4.5 also includes an XQuery editor, XQuery debugger and XQuery processor. Stylus Studio 4.5 includes integrated debugging with traceability for both J2EE and .NET environments. www.exln.com

TOPOLOGI LAUNCHES NEW XML/SGML EDITOR

9/30/2002

Topologi Pty Ltd. announced the Topologi Collaborative Markup Editor to support the lifecycle processing of large and complex XML/SGML documents, from initial conversion of unstructured source data to final validation and preview using commonly available typesetting and formatting tools. The Topologi Collaborative Markup Editor: Community Edition can be downloaded from Topologi's website for an evaluation period of up to 30 days. An introductory level registration fee for single users is US\$59. An annual site license for an unlimited number of users costs US\$5,000 including support. A high-end version of the editor will be released later in 2002. A plug-in to FrameMaker Server (and soon for other composition engines) is available to allow quality typesetting and PDF generation from a shared server. www.topologi.com

ANTENNA HOUSE UPDATES XSL FORMATTER

9/30/2002

Antenna House, Inc announced that an upgrade to its XSL-FO processor [XSL Formatter] to V2.3. V2.3 provides significant enhancements in layout function capability and multilingual formatting function capability, including implementing XSL float feature for page layout and UNICODE BIDI (bidirectionality) for mixing right-to-left and left-to-right languages. By using the XSL Formatter V2.3 PDF Option, it's possible to do layout the multilingual publications with the flexible mixture of Latin, Cyrillic, Greek alphabet, CJK (Chinese, Japanese, Korean), HAT (Hebrew, Arabic, Thai) and output to PDF. In V2.3, links and bookmarks are automatically created in PDF using Distiller. In addition, both EPS with preview image and EPS without preview image can be embedded in PDF via Distiller. By using a Plug-in (MathPlayerV1.0), the formula written by MathML can be embedded. www.antennahouse.com

DIVINE INTEGRATING CONTENT MANAGEMENT & COLLABORATION

9/26/2002

divine, inc. introduced a solution for automating the entire content lifecycle. Combining its enterprise content management, collaboration and search technologies, divine's integrated approach to managing the content lifecycle enables dynamic and timely collaboration not only as content is developed, approved, and published, but also when ideas are generated and refined and, ultimately, when key audiences interact with the content on Web sites and portals. With its integrated solutions approach, divine is helping firms address two key aspects of the content lifecycle: the idea-generation phase in which business users research and collaborate to develop and refine content; and the delivery phase when content, complemented by other technologies, is used to enable intelligent interactions through content-driven applications and business initiatives. Without effectively addressing these areas, companies will never be able to fully leverage content as a strategic corporate asset. www.divine.com

OPEN TEXT INTRODUCES LIVELINK DOORWAYS

9/25/2002

Open Text Corporation introduced Livelink Doorways, which allows users in Livelink to access and use content from other repositories within an enterprise. Livelink Doorways gives companies a content integration framework that enhances collaboration by bringing together all of a company's content, no matter where it resides, and making it available to users collaborating in Livelink. The first release of Livelink Doorways offers access to Documentum 4i and a variety of file systems, including Microsoft Windows, UNIX NFS and Oracle iFS. In later releases, Open Text expects to add connectors for Hummingbird DM and DOCS Open, Lotus Notes and Domino, and other repositories, as well as provide a means for partners and customers to develop connectors to their own unique repositories. With Livelink Doorways, other content sources are mounted as containers or "doorways" within users' Livelink workspaces. Folders, directory structures and documents from these content sources are navigable from each doorway within Livelink, using the source's native hierarchical structure and respecting each user's permissions to access such content. www.opentext.com

ADVENT 3B2 & DMSI TEAM

9/24/2002

Advent 3B2 Inc. has announced an agreement with Document Management Solutions Inc. (DMSi) to jointly market and provide services and solutions for automated document production in the United States of America. DMSi will have access to the 3B2 software and its key per-

sonnel will have extended training in the product. The two companies will co-operate in advancing system solutions for clients and providing service and support. www.3b2.com, www.dmsi-world.com

iPHRASE ANNOUNCES PARTNER PROGRAM FOR ISVs

9/24/2002

iPhrase Technologies, Inc. unveiled its partner program for independent software vendors (ISVs). Companies that wish to integrate a self-service search platform can now partner with iPhrase. iPhrase One Step adds value to ISVs in a variety of different markets, including: Content Management, Enterprise Portals, CRM, eCommerce, EAI, Databases, and Business Intelligence/Management Reporting. Partner members will receive access to the needed tools and program support necessary to develop their software integration. These tools include: Developer versions of the One Step platform to integrate and test applications; Advanced documentation, support and education to facilitate the integration process; A standardized certification process to test and validate the scope and quality of the software integration by the ISV developer; and Marketing support to help the ISV bring their iPhrase certified integration to the marketplace. www.iphphrase.com

VIGNETTE & EPICENTRIC ANNOUNCE PORTLET LIBRARY

9/23/2002

Vignette Corp. and Epicentric Inc. announced the availability of the Vignette V6 Portlet Library for Epicentric Foundation Server. The portlet library enables joint customers to integrate and access Vignette-managed content across multiple Web sites and portals through the user interface of the Epicentric Foundation Server. The Vignette V6 Portlet Library for Epicentric Foundation Server includes a set of content management and delivery portlets that facilitate content contribution, workflow, and site and channel management from Vignette V6 to the Epicentric portal. In addition, the portlet library leverages Epicentric Foundation Server security, customization and administration tools within Vignette V6. www.epicentric.com, www.vignette.com

SUN & AVALTUS IN ENTERPRISE LEARNING CONTENT MANAGEMENT PARTNERSHIP

9/23/2002

Sun Microsystems announced an agreement with Avaltus Inc., a provider of enterprise Learning Content Management Systems (LCMS), that enables customers to combine a standards-based learning management and learning content management solution using the Sun Enterprise Learning Platform and the Avaltus Jupiter LCMS Suite. This new solution also promotes both content and technology standards that can provide more flexibility to organizations in tailoring learning environments for their employees. The solution supports the Shareable Content Object Reference Model (SCORM), a set of interrelated technical specifications. These specifications enable the reuse of Web-based learning content across multiple environments and products by establishing the integration between learning management and content management systems. www.avaltus.com, <http://sun.com>

SIMSTAR ANNOUNCES ALLIANCE WITH INTERWOVEN

9/19/2002

SimStar Internet Solutions officially announced its alliance with Interwoven, Inc. The alliance, which provides SimStar with access to Interwoven resources, such as Interwoven TeamSite

software and Interwoven OpenTransform, is designed to strengthen the company's delivery capabilities for its behavior-based pharmaceutical e-marketing solutions. The collaboration will serve to enhance SimStar's processes for content management, workflow, knowledge management, and content conversion. www.simstar.com

CAMBRIDGEDOCS DISCLOSES STRATEGY

9/17/2002

CambridgeDocs (was XYZ Technologies, Inc.) disclosed details on the company's long-term product and architectural strategy, the XML Content Backbone. The company will focus on unstructured content such as that contained in Adobe PDF, Microsoft Word, plain text and HTML documents. The XML Content Backbone is a software platform that will integrate the unstructured content from disparate systems across the enterprise and from the extended enterprise. The XML Content Backbone will be able to migrate, integrate, route, and assemble content from document management systems, content management systems, groupware systems, desktop applications, and publishing systems using any industry XML standard such as DocBook, RIXML, HR-XML, NewsML, LegalXML or any custom application XML schema. The first CambridgeDocs product will address the need to convert Adobe PDF, Microsoft Word, and HTML documents into "meaningful" XML and where the meaning of XML is defined by the particular needs of the user. www.cambridgedocs.com

TRIDION RELEASES R5 PRODUCTS

9/17/2002

Tridion, has launched a range of products to help businesses control content from and across multiple channels. Tridion R5, the enterprise solution, comprises a series of offerings enabling companies to create, manage, distribute and deliver all forms of content to any target source. Tridion Client Connector enables users to create content with their preferred desktop application e.g. MS Office. Tridion Content Porter allows existing content to be imported and converted to XML, while Tridion Business Connector allows content from existing ERP, CRM or document management systems to be re-purposed. Tridion Content Manager manages the full lifecycle of the content. With in-depth use of XML and XML schemas it allows information to be assembled and re-used. It also incorporates multi-lingual capabilities. Tridion Content Distributor can dispatch content to any delivery platform. Tridion Web & Application Server Integration offers a delivery infrastructure to Web and application servers. Tridion Portal Server Integration offers integration with portal vendors, while Tridion Print Integration allows the production of customised print publications. The products are available from October 7th, 2002. www.tridion.com

ARBORTEXT FOCUSES ON LIFE SCIENCES

9/17/2002

Arbortext, Inc. announced a focus on the life sciences industry that includes new product capabilities and new applications. This initiative enables pharmaceutical companies, medical device manufacturers and biotechnology companies to bring new and revised products to market faster while complying with complex regulatory requirements. The new applications are a combination of Arbortext core products, partner integration and consulting services. Arbortext's Epic software provides a way for organizations to create, manage and publish complex business and technical documents, such as new drug applications, product information (package inserts, leaflets and labels), and written documentation of procedures that support the methods, facilities and controls used in the preparation, processing and packaging of pharmaceutical and medical products. Some of the new product capabilities to support this initiative include:

Change tracking, Enhanced API including support for Active-X, Stronger integration with Documentum's content management system, and Digital signatures and watermarks.
www.arbortext.com

FATWIRE SOFTWARE LAUNCHES UE STUDIO 2.0

9/16/2002

FatWire Software announced the release of UE Studio 2.0, the latest edition of its development tool to build and automate Web applications. An add-on to FatWire's enterprise-level dCM software UpdateEngine, UE Studio 2.0 offers expanded functionality to maximize developer efficiency, and ease the application backlog. UE Studio 2.0 provides tools to two types of users: programmers, who develop content centric applications and "power users," who understand and support the content management system but do not necessarily have programming skills. The new JSP (Java Server Page) Builder gives developers the ability to automatically generate JSP pages which displays the managed content within a framework that includes navigation as well as provides a foundation for laying out the look and feel of the site. The new Page Builder generates static HTML pages for solutions that do not require dynamic access to content, and the new Data Import Tool imports all types of structured content into the UpdateEngine managed content repository. www.fatwire.com

OPEN TEXT INTRODUCES CONTENT MANAGEMENT OFFERING FOR LIVELINK

9/16/2002

Open Text Corporation is extending Livelink to offer integrated Web content management. The new offering will tie content management into Livelink's knowledge management and collaborative capabilities. Information from other applications such as databases, document management repositories and other enterprise applications can be dynamically extracted and integrated into the website. Content Management for Livelink will also be fully integrated with Livelink's collaboration tools, enabling workflows to be used for content approval processes, and providing access to real-time collaboration with Livelink MeetingZone. Content Management for Livelink will include: Content authoring and creation, Multilingual site creation and management, Content approval and staging, Content distribution and syndication, and Content delivery and personalization. Content Management for Livelink will also provide out-of-the-box integration with application servers and personalization engines, such as BEA's Weblogic and IBM's Websphere. www.opentext.com/livelink

EKTRON RELEASES CMS100 VERSION 2.0

9/16/2002

Ektron, Inc. announced the release of Ektron CMS100 version 2.0. In addition to supporting Microsoft ASP and Macromedia ColdFusion, the new release supports Microsoft ASP.NET and PHP platforms. Ektron's browser-based content management solution provides core content authoring and publishing capabilities for US\$499. Ektron CMS100 offers a familiar word processor-like editing toolbar and intuitive interface for content publishing by non-technical users. Web professionals can easily configure and customize the solution to maintain control over navigation, look and feel, and other site infrastructure. Beyond added support for ASP.NET and PHP, Ektron CMS100 now offers user-controlled dynamic navigation, whereby non-technical content contributors create dynamically generated navigation based upon content in the database. Additional enhancements include improved style sheet support, new search functionality, a new interface for editor configuration, and improved branding capabilities. www.ektron.com

IPEDO INTRODUCES DYNAMIC INFORMATION SUITE VERSION 3.2

9/12/2002

Ipedo introduced the latest version of its Ipedo Dynamic Information Suite, featuring several product enhancements to support the growing use of XML in content-driven applications. The Ipedo Dynamic Information Suite Version 3.2 introduces: Virtual documents that can include references to any arbitrary pieces of content in other documents or in other systems through XML views; Full text search integrated with XQuery; Indexes that are automatically managed and optimized based on use and query patterns on individual XML documents and document collections; and support for user-defined hierarchies within document collections. The Ipedo Dynamic Information Suite Version 3.2 will be available starting in the fourth quarter of this year for Windows 2000, Windows NT, Sun Solaris and Red Hat Linux. Pricing is on a per-CPU basis. www.ipedo.com

X-HIVE & ARBORTEXT ANNOUNCE CO-MARKETING ALLIANCE

9/12/2002

X-Hive Corporation and Arbortext, Inc. announced they have signed a co-marketing agreement. X-Hive Corporation and Arbortext will work closely together to develop and market a seamless integration between X-Hive Corporation's native XML database, X-Hive/DB, and Arbortext's Epic software. This integration will enable X-Hive/DB users to edit and publish XML documents with Epic, and Epic users to take advantage of X-Hive/DB to manage XML documents. www.x-hive.com, www.arbortext.com

FAST SEARCH SEARCHES FLASH CONTENT

9/12/2002

Fast Search & Transfer (FAST) announced that Macromedia Flash content and applications can now be searched by users of its search technology showcase site www.alltheweb.com, and by FAST's worldwide portal partners who utilize FAST Web Search. www.fastsearch.com

ATOMZ ANNOUNCES ENHANCEMENTS

9/10/2002

Atomz announced numerous enhancements to its Atomz Publish and Atomz Search applications. Among the enhancements announced by Atomz is support for Macromedia Flash content, including Macromedia Flash MX, within Atomz Search. Atomz Search supports the indexing and crawling of Macromedia Flash content and applications. Other enhancements announced by Atomz include: Advanced Phrase, Acronym and Synonym Support (Atomz Search); Advanced Template Management (Atomz Publish). Related Content Display (Atomz Publish and Atomz Search); Forms-Based Authentication (Atomz Search); Advanced Search Templates Management (Atomz Publish and Atomz Search); Time-To-Search Display (Atomz Search); and Media Manager (Atomz Publish). www.atomz.com

GREY ZONE RELEASES SECUREZONE 5 FOR XSERVE & MAC OS X

9/10/2002

Grey Zone Inc. announced the immediate availability of SecureZone 5 for Apple's Xserve and Mac OS X Server version 10.2. Grey Zone's solution offers Xserve users essential features required of Extranets, such as access control, content management, personalization, and presen-

tation. Grey Zone's software is an end-to-end solution that provides a foundation for extranets as well as corporate intranets and public web sites. www.greyzone.com

EMC & DOCUMENTUM ANNOUNCE CENTERA AND DOCUMENTUM INTEGRATION

9/10/2002

EMC Corporation and Documentum announced they have completed integration between Documentum's ECM platform and the EMC Centera Content Addressed Storage (CAS) solution. The two companies also announced plans to provide Documentum 5 with extended fixed content management (FCM) features based on Centera. EMC Centera is an online storage architecture specifically designed to address the unique storage requirements of fixed content, such as photos, videos, audio, graphics and web content. Centera integration with Documentum's ECM platform provides a total solution for managing, distributing, exchanging and storing large volumes of content across an entire enterprise. www.EMC.com, www.documentum.com

DOCUMENTUM ANNOUNCES DOCUMENTUM 5

9/10/2002

Documentum announced Documentum 5, the latest version of its ECM platform. The Documentum 5 platform offers: Enhanced usability and deployability to enable widespread adoption of content management; Expanded enterprise collaboration that can be extended to any business process; Fixed content management, to manage images, reports and records; Advanced trust, security and compliance services to provide protection of content assets; and a standards-based development environment for easy application development. Documentum 5 offers unified content services for enterprise document management (EDM), web content management (WCM), digital asset management (DAM) and now fixed content management (FCM) — records, reports and scanned images — in a single, integrated platform. www.documentum.com

WEBWARE ANNOUNCES PRODUCT LAUNCH SOLUTION

9/10/2002

WebWare Corporation announced the introduction of the WebWare Product Launch Solution. Based on WebWare MAMBO digital asset management technology, the Product Launch Solution is an outsourced service hosted by WebWare that reduces the time it takes for a company to bring a product to market. The solution is completely web-based. The Product Launch Solution enables secure digital distribution over the Internet of both in-process and ready-to-use marketing collateral to global recipients in advance of a launch. Such items might include video and audio content, 3-D graphics and compound documents such as data sheets, sales presentations, marketing white papers, official logos, product shots, print advertisements (in multiple sizes and formats), product messaging, and marketing support materials. The solution features: Cross-media support for managing multiple media and production file types; Simple file upload and cataloging process; High-availability, enterprise security; full back-up; and Rapid ramp-up with minimal organizational distraction. This is the first in a series of WebWare Business Solutions planned by the company. Pricing and availability will be announced shortly. www.webwarecorp.com

CROWNPEAK RELEASES PUBLISHER ADVANTAGE CMS

9/9/2002

CrownPeak Technology announced the release of Publisher Advantage, a special publishing industry edition of Advantage CMS, its ASP-based content management solution. Developed in response to the specific needs of the publishing industry, this special version of Advantage CMS offers a series of enhanced capabilities: greater user manageability and control over content and digital assets, infinitely configurable workflow, improved proofing tools, integration with ad serving components, e-mail newsletter creation and management, performance monitoring, increased report management capabilities, and an exclusive "content similarity engine." In addition, Publisher Advantage delivers enhancements in the areas of directory management, e-mail marketing and ad serving. www.crownpeak.com

EASYPRESS ANNOUNCES XML/QUARK ROUNDTRIP

9/9/2002

Easypress Technologies announced Atomik Roundtrip 1.0 — software to provide full, bi-directional XML support for QuarkXPress versions 4 and 5. Atomik Roundtrip will enable users to import XML into QuarkXPress and faithfully re-export it. With a single click, users can update the QuarkXPress document if the source XML document(s) change and update the source XML document(s) if the QuarkXPress document changes. XML content can be added to QuarkXPress documents through a drag and drop process or by using the Atomik Roundtrip Placeholder technology to import individual XML elements or collections of XML elements into predefined templates. A Documentum integrated version of Atomik Roundtrip will be available during the fourth Quarter of 2002. Direct support for other content management systems will follow. Product is expected to ship at the beginning of October 2002. The suggested retail price for a single user licence of Atomik Roundtrip 1.0 will be £3,495, \$4,495 or €5,662. The suggested retail price for a 10-user licence of Atomik Roundtrip 1.0 will be £9,950, \$13,500 or €16,119. www.easypress.com

ARTESIA TECHNOLOGIES LAUNCHES ASSET-CENTRIC WORKFLOW

9/9/2002

Artesia Technologies unveiled TEAMS, version 4.4. The new version's enhancements for collaboration and workflow are designed to redefine the definition of an asset through the inclusion of new information capturing the entire lifecycle of the asset. The core features of TEAMS 4.4 are newly integrated Asset-centric workflow capabilities that enhance the collaborative aspects of Artesias enterprise architecture for managing digital assets. This includes the ability to immediately determine where an asset has been used throughout its history, who has participated in its usage, review or licensing, and the nature of these projects or processes that have utilized the asset in one form or another. Organizations are able to eliminate the need for the separate, proprietary workflow tools associated with each production processes. By embedding many aspects of workflow within the asset itself, where it resides alongside other attributes governing format, rights and permissions, and other business information, TEAMS 4.4 can ensure that this information is readily available and actionable not only to the specific user and their co-workers. www.artesia.com

XEROX INTRODUCES DOCUSHARE 3.0

9/5/2002

Xerox Corporation introduced DocuShare 3.0, its Web-based document and content management software. Incorporating features that meet the needs of both small businesses and global

enterprises, the new software is designed to deliver performance and simplicity starting at less than \$4,500. DocuShare 3.0 provides a complete document management solution fostering better office collaboration among workers in document-intensive environments. Built on an all-new Java platform, DocuShare 3.0 runs on Windows, Linux and Solaris systems. The software will be sold through Xerox direct sales representatives, agents, concessionaires, Xerox Business Partners and Teleweb sales channels. The entry-level list price for a complete DocuShare system with 10 seats is \$4,145, and a 100-seat system is \$9,995. Worldwide availability begins Sept. 30. www.xerox.com/docushare

EBXML MESSAGING SERVICE SPECIFICATION APPROVED AS OASIS STANDARD

9/5/2002

ebXML Messaging Service Specification version 2.0 has become the newest OASIS Standard, completing a recent election by the consortium's membership at-large. The ebXML Messaging Service standard, which provides a secure method for exchanging electronic business transactions using the Internet, carries forward work initiated by OASIS and the United Nations Centre for Trade Facilitation and Electronic Business (UN/CEFACT). To attain status as an OASIS Standard, ebXML Messaging Service v2 was first approved by its development team as an OASIS Committee Specification. After being implemented by a minimum of three organizations, it then underwent a 90-day open review, before the final balloting of OASIS members. ebXML Messaging is one of a suite of specifications that enables enterprises of any size and in any geographical location to conduct business over the Internet. www.oasis-open.org

SOFTWARE AG & ACTIVE DATA EXCHANGE TO PROVIDE ENTERPRISES CONTENT DISTRIBUTION

9/4/2002

Software AG, Inc. and Active Data Exchange, Inc. announced a strategic sales and technology alliance to provide a content distribution management solution across multiple enterprises. Active Data Exchange's content distribution management solution, Active Data Syndicator version 4, embeds Software AG's Tamino XML server. The result of the Software AG - Active Data Exchange alliance is an end-to-end solution that enables companies to create new revenue streams by taking advantage of existing content. Software AG's Tamino XML server and Active Data Syndicator both store, process and deliver XML. The time required to aggregate and transform existing content is decreased, resulting in faster distribution of vital information to internal and external customers. Software AG has agreed to distribute Active Data Syndicator in North America. www.softwareagusa.com, www.activedataX.com

PLUMTREE RECEIVES UNSOLICITED OFFER

9/4/2002

Plumtree Software, Inc. announced that it has received an unsolicited conditional offer from the Sutter Opportunity Fund 2, LLC to acquire all outstanding shares of Plumtree common stock in exchange for \$2.00 in cash and a promissory note due in five years. The offer came in the form of a letter dated September 3, 2002 from Sutter Capital Management of San Francisco to Plumtree CEO John Kunze. Plumtree's Board of Directors will evaluate the terms of the offer in due course. Prior to receiving the letter, Plumtree had no contact with Sutter Capital Management. www.plumtree.com

TOPICALNET ACQUIRES LIGHTSPEED INTERACTIVE

9/3/2002

TopicalNet, Inc. has agreed to purchase substantially all of the assets of Lightspeed Interactive, Inc. Financial terms of the acquisition were not disclosed. The new company will be named LightSpeed Software, Inc. Lightspeed Interactive is the fifth company TopicalNet has acquired since April 2001 in its strategy to assemble a complete, automated content acquisition, content management, collaboration and delivery product for mid-market customers. It started by acquiring Internet Profiles Corporation (I/PRO) in April 2001 for its online audit and analytics capabilities. In October of 2001 TopicalNet purchased Collectively Sharper and its content integration technology, followed in December by Teralytics and its tools dedicated to analyzing customer interaction data. In June of 2002, TopicalNet purchased Wego Systems and its Ready-Portal portal development and collaboration product. By combining Lightspeed Interactive's content management products with TopicalNet's XML based Search, Classification/Metadata, Portal/Collaboration and Analytics technologies, customers manage both internal and external content from research and acquisition through the creation, delivery, collaboration and repurposing process. The corporate headquarters will be relocated from Woburn, Mass. to San Francisco. www.TopicalNet.com

OBTREE & BEA PARTNER

9/3/2002

Obtree Technologies Inc. and BEA Systems Inc. announced the integration of Obtree's content management solutions with the BEA WebLogic Platform 7.0. Under the terms of agreement, the two companies have combined their technologies, allowing developers to integrate Obtree's personalized content on the BEA WebLogic Platform. Obtree's integration into the BEA WebLogic Platform delivers content dynamically into the portal or application framework that corresponds to the rights and interests of a particular user. Rather than store content in a repository for delivery to the portal according to a predefined framework, Obtree personalizes content delivery to fit user profiles. Obtree enables users to edit or update the content provided them in real-time through the portal interface. www.bea.com, www.obtree.com

AGARI TARGETS PUBLISHING WITH MEDIA STAR 2.2

9/3/2002

Agari Mediaware, Inc. announced the introduction to the publishing industry of Media Star 2.2. Agari's Media Star product suite allows the integration of publishing applications, non-intrusively, without making changes to existing applications or content. With it publishers can perform federated search, retrieval, and processing of content across geographically dispersed applications, link distributed content such as layouts, photos, artwork, pricing, rights, sales, and distribution data, and integrate systems acquired through mergers and acquisitions. The Agari product suite provides a solution for automating content value chain business processes (content creation, transformation, management, distribution, and consumption). Agari's Enterprise Business Asset System connects content that is often dispersed across departments, business units, and partners so that it can be used across the enterprise. The metadata can be transformed between various formats on the fly, eliminating the need for centralized, common metadata schemes. www.agari.com

SUBSCRIPTION FORM

You can also order on our **secure** website www.gilbane.com.

Please start my electronic subscription to the Gilbane Report for **only \$99**. (10 issues/year). Subscription includes access to HTML and PDF versions at www.gilbane.com. (Call for print subscriptions, site license prices, and back issues.)

I am eligible for an affiliate discount* _____ Affiliate organization _____ Tracking # _____

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

Name as on card: _____ Number _____
Signature _____ Expiration date _____

Name _____ Title _____
Company _____ Department _____
Address _____
City _____ State/Province _____ Zip/Postal Code _____
Country _____ Tel. _____ Fax _____ E-mail _____

Checks from outside the U.S. should be made payable in U.S. dollars.

Funds may be transferred directly to our bank, please call for details.

Mail this form to: Bluebill Advisors, Inc. 763 Massachusetts Ave., Cambridge, MA 02139, USA.

You can also place your order at www.gilbane.com or by phone (+617.497.9443), or fax (+617.497.5256).

CALENDAR

Integration 2002 – Forum XML & Web Services. November 13-14, 2002, *Palais des Congrès – Paris, France*. TechnoForum's 5th annual conference on XML and integration. Among the topics covered this year : Web Services, Enterprise Application Integration (EAI), Corporate Portals architectures, Content management, XML-EDI and ebXML, XML standards, Supply-Chain & B2B Integration, .Net & J2EE Architectures, and XML databases. www.gilbane.com/events/programme_integration_2002.pdf, www.technoforum.fr/index.html

XML 2002. December 8-13, 2002, *Baltimore Convention Center, Baltimore, MD*. The XML Conference & Exposition 2002 is the largest and longest-running annual gathering of XML users and developers in the world. This event is well known in the XML community for attracting high quality speakers and attendees. **Special Offer to Gilbane Report subscribers: Save \$300 off the cost of a Conference Gold Pass.** Login to the Gilbane subscribers section (www.gilbane.com) to get the discount priority code to use on the registration form. (Discounts cannot be combined.) <http://www.xmlconference.org/xmlusa/>

Documation France 2003. March 17-19, 2003, *CNIT La Défense – Paris, France*. Our 10th annual Documation France with TechnoForum covers content management, enterprise portals, and electronic document technologies. Mark your calendars and stay tuned for more information. <http://www.gilbane.com/documation03.html>, www.technoforum.fr/index.html

See www.gilbane.com/events.html for more events and updates.

© 2002 Bluebill Advisors, Inc. - all rights reserved. No material in this publication may be reproduced without written permission. To request reprints or permission to distribute call +617 497.9443. The Gilbane Report is a registered trademark of Bluebill Advisors, Inc. Product, technology, and service names are trademarks or service names of their respective owners.

The Gilbane Report is published 10 times a year (monthly, except for August and December). The Gilbane Report is an independent publication offering objective analysis of technology and business issues. The Report does not provide advertising, product reviews, testing, or vendor recommendations. We do discuss product technology that is appropriate to the topic under analysis, and welcome product information from vendors. Letters to the editor are encouraged. Send to: editor@gilbane.com. Visit our web site at www.gilbane.com

ISSN 1067-8719