

THE GILBANE REPORT™

Content, Computing, and Commerce – Technology & Trends

Vol. 10, No. 3
April, 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

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

Subscriptions:
sales@gilbane.com
(617) 497.9443

Customer Service:
customersevice@gilbane.com

Consulting Inquiries:
Steve Paxhia
steve@stratadatapartners.com
(617) 868.4668

Gilbane Report White Papers:
Bill Trippe
bill@gilbane.com
(617) 497.9443

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

IN SEARCH OF SEARCH SOLUTIONS

Search, and search-related, technology is enjoying renewed interest these days. There are dozens of products, analysts are busy selling reports, and businesses are realizing they could benefit from even incrementally better search capabilities. It is time to take a look at what's available.

We are constantly being told, as if we needed reminding, that we are so overwhelmed with information we often can't find the specific knowledge nuggets we need. The problem is not just the *volume* of information, but also the *variety* of information types, and the lack of information *organization*. Much of the current wave of development is focused either on the variety problem (structured, unstructured, and rich media) or on the organization problem (categorizing and taxonomy tools). Indeed, vendors often differentiate themselves based on which of these they specialize in. Businesses however, need to look at all three aspects of the problem. Organized information is both easier to find and more useful when (re-)organized for specific uses once found. While there are many situations where a Google-like search is just what you need, many business applications require at least the ability to store what has been found for further use without having to recreate imaginative search queries. Organization (of which categorization is one aspect) and search should be considered together when building IT strategies.

This month, Sebastian provides you with a way to get started by laying out a high-level taxonomy of the market, and some guidance on what to think about as you consider investing in new search technologies.

CONTENTS

In Search of Search Solutions	2
Industry News	13
The Gilbane Conference on Content Management @ Seybold	26
Subscription Form & Calendar	28

IN SEARCH OF SEARCH SOLUTIONS

If you can't find it – it isn't there.

Prediction: You are often frustrated and ultimately fail when looking for information. Sometimes this happens on the Internet and sometimes you are working within a specific application.

Do I have ESP? Is this an example of some sort of personalization where each Gilbane Report is customized just for you? Sadly, it is neither of these. The simple truth is that if you use a computer, search technology has let you down on more than one occasion. There are numerous reasons for this, and not all of the blame can be placed upon search technology. In fact, intelligent search technology is deceptively complex. This month, we will look at the market dynamics that are pushing search technology to new heights, categorize the various flavors of search technology that have emerged, and review a cross section of the software vendors that are hoping to solve your search problems once and for all.

MARKET DRIVERS

Our use of technology in the workplace and at home continues to evolve rapidly. Each new application, device, and media type brings with it new technology and usability requirements that the search technology of only a few years ago simply cannot support. The result is that there has been a revival for search technology vendors. The major factors driving this growth in search technology include the following.

Explosive increase in the *volume* of content

- **Raw content:** It has been estimated that the total volume of information on the planet is doubling every three years. Much of this content is in digital form.
- **Published content:** The volume of content becoming available publicly (on the Internet) is multiplying at an even greater rate.
- **Connected content:** As applications become integrated across broader sections of business and society, each individual application has access to greater amounts of content that had previously been contained within one or more “stove pipe” applications.

Increased *variety* of content

- **New Media Types:** Rich media formats including video, audio and images for traditional distribution and emerging digital channels have completely changed the rules of the search game.
- **Metadata:** The increased use of metadata (data about content) to capture rights, usage, ownership, *etc.* in both general and industry-specific ways has radically altered how search criteria needs to be applied. Often metadata is stored and managed independently from the content it describes. This alters not only the search algorithms, but also how results need to be returned and managed.
- **Structured Content:** The rapid adoption of the XML family to capture domain semantics (meaning for particular uses), presentation rules as well as

metadata has resulted in further requirements for search algorithms, result management and presentation.

Greater variety and numbers of users

- **Uses and roles:** The proliferation of the Internet into virtually every facet of our daily lives from daycare to automobile shopping has brought with it an equal number of new use cases and categories of users. The kinds of searches, user expectations on accuracy, and completeness and assumptions about privacy and reuse often change quite dramatically across these new user profiles.
- **Skill levels:** As web-based applications find themselves in increasingly specialized uses, the skill levels of users are also becoming increasingly diverse. Novice, elderly, toddler and special needs users all bring unique requirements, as do domain experts such as medical doctors, lawyers, chemists and engineers, all of whom have unique search use cases and varying degrees of willingness and ability to be trained in specialized search techniques.
- **Languages:** Global communication and information sharing as well as increased access to the Internet play havoc with information management in general and search techniques and their underlying assumptions in particular. Sorting, indexing and organizing content expressed in multiple languages requires special technical and operational considerations that a single language environment simply does not need to consider.

WHAT DOES SEARCHING INCLUDE?

Searching is the process of matching a user's request with a set of results that meet that specific request. Beyond that, the definition gets very complex very quickly. Database, document management, video archives, digital libraries, web sites and every other kind of information store act as incubators for evolving search technologies. As the Internet infrastructure connects all of these various data stores, there will be increasing pressure to not only improve search technologies but also to integrate them. The following overview is intended to provide a holistic overview of basic search functionality that can be used to evaluate the diverse array of search products in the market today. The objective here is to optimize for accuracy while clearly skimping on precision.

Ingestion

In order to provide the response times that users have come to expect, search engines must build indexes and collect statistics on the content in before it can process its first query. The following are functions that are often, although not always, present in search engines.

Content extraction functions include:

- Filters that look inside content to extract information that are ordinarily hidden within a proprietary format such as PowerPoint, PDF, *etc.*
- Loggers analyze, deconstruct and extract information from time-based content such as video and audio.
- Parsers analyze structure and extract information and raw content from structured content such as XML.

- Transformers and encoders generate proxy and alternate versions of content for simplified searching and previewing of content.

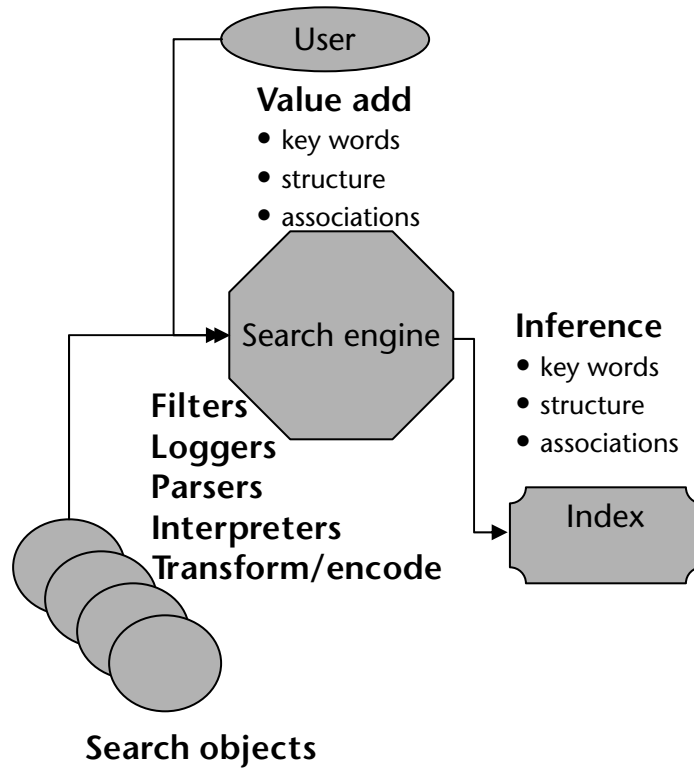


Figure 1. Ingestion and Indexing

Content organization and categorization functions include:

- **Key words:** These are terms that users would associate with content, *e.g.*, news, celebrities, software, *etc.*
- **Structure:** Semantic (meaning) and formatting structure can be captured, *e.g.* owner *versus* page numbers.
- **Associations:** This is a very broad category that includes relationships (versions), topic maps (states to cities), linguistic (monkey is an animal), *etc.*

These functions can be provided by users/administrators but are increasingly being offered as intelligent inference services within the search engine.

The shift from manual to automatic ingestion is at the heart of much of the innovation in today's search technology. The cost in person-power and time required to manually tag content with key words and associations often preclude the use of search technology from all but the most critical applications. As automatic ingestion becomes increasingly accurate and sophisticated, the market for advanced search technology can be expected to explode.

Query Processing

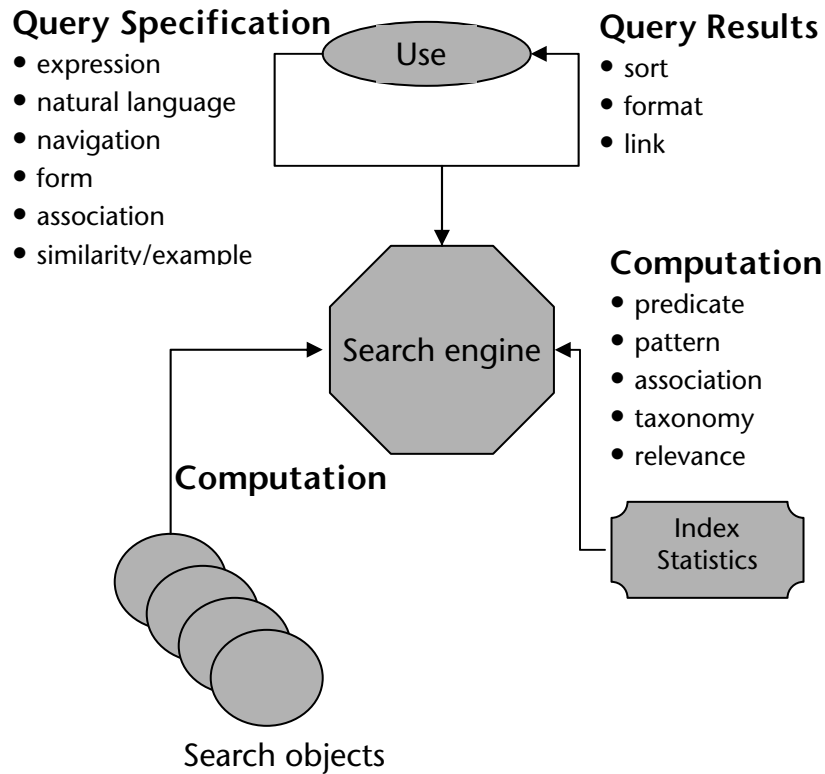


Figure 2. Query Processing

Once content has been analyzed and an index has been built, searches can now be submitted, processed and results returned. This section describes functions that are often although not always present in search engines.

Users must articulate their request. SQL queries, natural language interpreters, navigation *a la* Yahoo!, form-based, inference, and query by example (including cut and paste of images, audio and video) are all available to users today.

At this point, the search engine takes over and starts crunching statistics and navigating various links. Expressions are resolved, patterns are matched, categories and associations are traversed and the resulting answer set is often sorted by content value (date, last name) or by relevance (proximity or *nearness* of match).

The resulting matches are ordered, formatted and presented to the user for inspection and to provide direct access to specific content referenced in the result set and *Voilà!*

In an ideal world, the user is then presented with a result set that includes:

- Meaningful references, *e.g.*, the summary or previews of each match accurately represents the actual match. Examples include thumbnails of a PDF page, a proxy of a video clip, or a gist of a piece of text.
- No false positives, *e.g.*, each result returned satisfies the user's search criteria.

-
- No false negatives, *e.g.*, there is no content that has been excluded from the search set that the user would have been interested to review.
 - Accurate sorting and relevancy ranking, *e.g.*, of the twenty thousand potential matches to a query, the first 20 results formatted on the first page truly are the most relevant.

Of course, in the real world users are often imprecise in the expression of their search criteria, search engines are limited in the variety of searches they can perform, and content is often poorly categorized or completely unavailable to the search application. This is the problem that enterprising search technology vendors are trying to solve today.

Some commercial solutions focus on the *automation* of categorization, metadata creation, and keyword generation. One challenge is how to generate this information in a precise enough way to map into the distinct and often contradictory models across industries and use cases. Academic institutions, intelligence agencies and corporations see the world through very different lenses and therefore expect content to be organized accordingly.

Another challenge is that different media types require very different technology to peer into and analyze content. The algorithms required to extract the semantic meaning from an audio track are quite distinct from those used to analyze a novel.

This is a seemingly impossible hurdle to clear in a completely general way for all use cases, but vendors are able to deliver impressive results by reducing the variety of content and contexts to be supported at any given point in time. Increasingly, cost savings and productivity enhancements are clearly validating this approach of increased capability over specific classes or categories of content.

Issues to be wary of include:

- Automatically generated metadata and categories are not likely to match industry standards that your organization and your trading partners may be considering, *e.g.*, PRISM, ICE, *etc.*
- While content may have many uses and be of interest to many different user communities, the metadata and categories are not likely to have the same transferability.
- If the ingestion components of a search engine cannot process specific media types, significant amounts of content may be inaccessible.

One approach to compensate for divergence and distributed indexes, categories and locations is the mixed search.

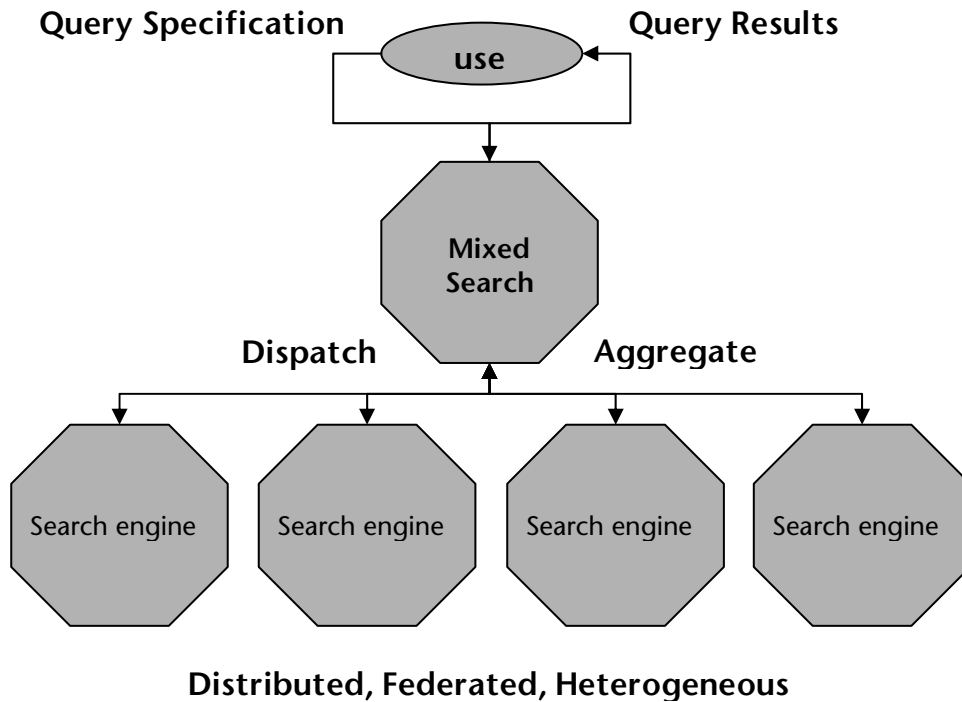


Figure 3. Mixed Search

A mixed search engine is one that accepts a search request and then dispatches localized versions of that search to multiple search engines. The results are then normalized, aggregated and returned as a single result set. Dependence on multiple search strategies that are often supplied by third party search technology can lead to some unexpected results, but it is the best and only way to search across all data regardless of location, format, or use. Mixed search solutions can be homegrown, be part of a content or digital asset management system, or be sold by search vendors.

CLASSIFYING SEARCH TECHNOLOGIES

There is no individual search technology that can cover the full spectrum of content types and search algorithms nor is there (yet) a single company that has integrated the search techniques to provide a one-stop-shop. It is therefore important to consider the strengths and weaknesses of each approach. The following framework is intended to provide some degree of order when evaluating the myriad of search products and technologies that are currently available.

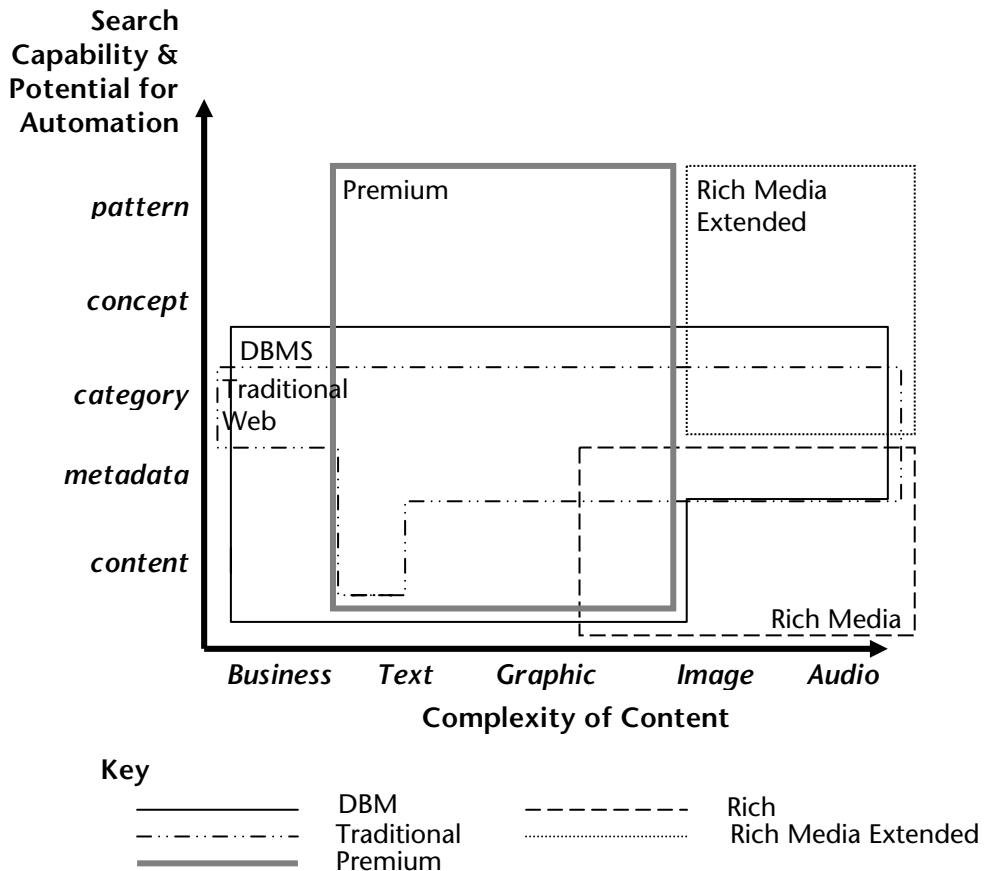


Figure 4: Classifying Search Technologies

Figure 4 illustrates how four families of search categories have covered the entire spectrum of searches. As we review vendor examples of each category, there are three important points to keep in mind.

- Individual products and the companies that offer them are constantly expanding their vision and scope. As such, this diagram is not intended to imply permanent limitations, rather to emphasize centers of excellence and historical success.
- The companies mentioned are not intended to provide an exhaustive list of search companies and their products. These represent a subset taken from a larger group that the author was quickly able to identify. This is a very crowded field.
- The author has not personally evaluated each product offering and cannot therefore warrant any product's quality or suitability for a particular purpose.

DBMS

Search technology in DBMSs has developed from query optimization engines that crunched out specific search strategies for relational queries to hybrid search engines that include category and cross media search. The most advanced of which now include references to content outside of the DBMS in question. This is part of a larger trend of the large DBMS provider's attempt to flip the content world upside down by making the DBMS the file manager for the enterprise

(rather than have the DBMS sit in the file system. As long as this paradigm is not uncomfortable and there is no issue with running all content through a single DBMS product, this technology can be quite comprehensive.

Traditional Web

Web-oriented search technology was delivered primarily through public portals (*e.g.*, AltaVista, Excite, Lycos, HotBot and, recently, Google). These products were relatively unsophisticated in terms of the search algorithms: they build indexes of all significant words and use those to look up documents. The quality of results depends on the range, sophistication and frequency of updating page links (the “crawling” process), and the differences in the algorithms used to rank the relevance of the results returned.

Premium Search

Premium search solutions are those that are based upon advanced search technology and are focused primarily on large, strategic collections of content and whose customers are willing to pay top dollar to provide the most robust search capabilities possible.

In 2002, the state of the art typically includes some combination of statistical, semantic, syntactic, and contextual methods to understand key concepts for the organization, enhancement, and utilization of relevant information. In other words, solutions use complex algorithms that use every possible computational method available to replace the need to manually define taxonomies, keywords, and metadata, and to properly assign content to these categories. Most of these algorithms will not be understood by business or IT managers, and it is difficult to determine how well they work. However, there are some very impressive demonstrations available and there is a lot of serious development going on. (We may delve further into the different algorithm types in an upcoming article.)

Rich Media

This typically includes specific fluency in logging and indexing time-based media (video and audio) to extract and synchronize clips, closed captioning, speech to text and other data extraction utilities.

Rich Media Extended

Extended rich media search includes advanced pattern matching across individual images as well as across time-based sequences. Facial image recognition and action recognition, *e.g.*, the scoring of a goal in sports, are both examples of this expanded media search capability.

Mixed Search

This is a pluggable architecture that includes the dispatch, aggregation and normalization of results across multiple, heterogeneous search engines.

	Traditional Web	DBMS	Premium Search	Rich Media	Rich Media Extended	Mixed Search ¹
Albert						
Altavista						
Answerfriend						
Applied Semantics						
Autonomy						
ClearForest						
Convera ²						
DreMedia ³						
eVision						
Fast Search & Transfer						
Fast-Talk						
Google						
H5						
IBM						
Inktomi						
Insightful						
InXight						
Iphrase						
LingoMotors						

¹ Mixed search solutions, as we have defined them, are mostly developed as in-house solutions, or are part of a solution from a vendor of enterprise software (content management and digital asset management especially) that works with multiple search vendors. We are focusing on the search vendors here. Over time you can expect to see some of these working as partners or by acquisition to provide mixed solutions.

² Convera has two products that combined provide this search "footprint." However, Retrievalware and Screening Room are not at this time seamlessly integrated.

³ DreMedia OEMs Autonomy's linguistic search with advanced video and audio search techniques giving it the largest search "footprint" in this chart.

LTU Technologies						
Mohomine						
Mondosoft						
Oracle						
Primus						
Quiver						
Sageware						
Semio						
Smartlogik						
Stratify						
Unifind						
Verity						
Virage						
Wherewithal						

Table 1. Mapping selected search vendors to the search techniques graph in Figure 4.

OBSERVATIONS

- The first obvious conclusion is that this is a crowded market.
- The majority of activity is currently focused on automating the ingestion, indexing and categorization of content. Both the Premium Search and the Rich Media Extended technologies invest heavily in automating the definition and population of information models. The reasonable premise behind this is that if it is too expensive and time consuming to organize and mark-up content, the vast majority of that content will never become searchable. While this is in fact true, it is also the case that auto-generated metadata and taxonomy models cannot be relied upon to facilitate the interchange of valuable content between organizations or to optimally preserve digital content for very long periods of time across multiple uses. Automating search ingestion and taxonomy generation is certainly a market-widening approach, but it can never fully displace careful and deliberate information modeling and content archiving.
- Web search is rapidly becoming a low cost commodity.
- Much of the cutting edge work is being done in the rich and extended rich media space. While it is not typical form in these articles, here is some homework for the interested reader. Visit www.dremedia.com. One of the features of this search technology is that one can edit the

xml-tagged textual transcripts of video. The index is used to automatically edit the original video to correspond with the cut and pasted text – this is a poor-man’s desktop broadcasting requiring no new editing skills beyond what one needs to edit email today. If the software works as advertised, this could get very interesting.

- The Premium Search is the most popular category in this high-level view of the market. In fact, there are a number of differences between some of these vendors and some would prefer to be thought of as offering categorization products rather than search. We’ll sub-divide this category in a future issue.

CONCLUSIONS AND RECOMMENDATIONS

There is a “principle of least surprise” that is based upon the premise that predictable software holds more value than software whose performance might occasionally be spectacular but cannot be relied upon to provide consistent and expected results. For all but the most dogged researcher that has plenty of time on their hands and a great deal of research expertise of their own, the principle of least surprise should probably sit at or near the top of the priority list when selecting and deploying search technologies.

Know the strengths and weaknesses of the search technology you use. If blind spots and particular strengths are well understood, then users can appropriately compensate and can assess the likelihood that they may have “false positives” in their result set or may have left “false negatives” behind. Don’t count solely on your ability to understand the relative merits of the sophisticated algorithms and linguistic, mathematical, and statistical theories they are based on. You need to test with your own content, including content from repositories you do not control, but need access to.

Assess user expectations and know how much search is enough. Causal and basic search requirements require only inexpensive and simple search tools. Advanced search requirements require a more careful analysis of user expectations, user expertise, the state of the content being searched (existence of metadata, *etc.*), and the suitability of available search technology.

Begin good housekeeping practices immediately to be best prepared to take advantage of emerging technologies. Wherever possible, capture as much descriptive metadata that may serve as useful search criteria in the future, *e.g.*, creation dates, authors, subject matter, rights and permissions *etc.* Much of this information simply cannot be inferred and must therefore be captured somewhere. Develop and utilize categorization and archival “best practices.” There are emerging best practices for digital preservation and that library science is assimilating the XML family, making living digital archives a reality. This should greatly simplify search requirements for large organizations and leave every category of information-intensive organization best prepared to take advantage of the many interesting and powerful search technologies that continue to emerge and mature.

Sebastian Holst
sebastian@gilbane.com

INDUSTRY NEWS

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

STELLENT ANNOUNCES SUPPORT FOR ORACLE9I APPLICATION SERVER

3/27/2002

Stellent, Inc. announced the Stellent Content Management system now supports the Oracle9i Application Server (Oracle9iAS). Organizations can use the Stellent Content Integration (CI) Kit to integrate business and Web content with other enterprise applications running on Oracle9iAS, giving users dynamic delivery of content, access to content management functionality and the ability to transform enterprise portals into content-rich business communities. Stellent allows Oracle9i e-business applications to access the business and Web content managed by Stellent, as well as Stellent Content Management functions such as full-text and metadata searching, library services, and workflow. www.stellent.com

ORACLE ANNOUNCES SUPPORT FOR ADOBE GOLIVE & ALTERCAST

3/27/2002

Oracle Corp. announced that Oracle9iAS enables full interoperability and enhanced performance between Adobe GoLive 6.0, a Web authoring tool, and Adobe AlterCast, a dynamic imaging server. By leveraging Oracle9iAS, Adobe GoLive and Adobe AlterCast software quickly and easily interoperate, enabling real-time development and accelerated delivery of content. Together, the three products allow Web designers and publishers to integrate and manage visually rich content with the dynamically generated JavaServer Pages (JSP) created by Java developers. In turn, Java developers can collaborate with Web designers and publishers to programmatically reformat and update visually rich images. Oracle9i Application Server supports J2EE 1.3, Web Services, SOAP, WSDL, UDDI, RosettaNet 1.1/2.0, ebXML, WebDAV, LDAP v3, SSL v3, and XML. www.oracle.com

INTERWOVEN ANNOUNCES ACQUISITION OF XYZFIND TECHNOLOGY

3/27/2002

Interwoven, Inc. announced the launch of its Enterprise Discovery Framework and has entered into a definitive agreement to purchase assets of XYZFind Corporation. Interwoven's Content Discovery Framework is an architectural layer that leverages metadata provided by Interwoven MetaTagger 3.0 and XML technologies provided by XYZFind to increase the efficiency and accuracy of information retrieval and reuse of enterprise content. Within the Discovery Framework, the XYZFind technology enables enterprises to organize XML content for easy access, quickly identify content elements to reduce redundancy, and promote efficient content retrieval. Interwoven's Content Discovery Framework allows companies to access, tag, store and retrieve enterprise content regardless of its format or where it is located. www.interwoven.com/company/features/content_discovery/index.html

UNISITE ANNOUNCES LAUNCH OF UNISITE CONTENT MANAGER

3/27/2002

UniSite Software announced the launch of UniSite Content Manager, an advanced Web content management solution that "can be implemented in as little as two weeks". The technology

transforms an existing Web site into a database-driven site supporting content for multiple channels-including wireless-and allows team members to collaborate and update content easily from anywhere in the world. For marketing professionals and Webmasters who want fast, efficient up-keep of their Web sites, UniSite offers advantages in an affordable, rapidly implemented system. UniSite Content Manager's unique technology scans an existing Web site, identifying and tagging content that needs to be managed, and structures it into a database-driven content management platform. Once a site is tagged, UniSite Content Manager is accessed through a Web browser while a workflow engine speeds updates and changes. UniSite can also manage multiple language versions of a site within the same structure.

www.UniSite.com

GAUSS ANNOUNCES SUPPORT FOR ORACLE9i APPLICATION SERVER

3/27/2002

Gauss announced that its ECM Suite, VIP Enterprise, now runs on Oracle9i Application Server (Oracle9iAS). The combination of VIP Enterprise and Oracle9i Application Server helps reduce the time and cost of distributing content to the Web for enterprises with decentralized e-business environments. Oracle9iAS provides the infrastructure to deploy distributed Web applications. VIP Enterprise provides content management capabilities for the global enterprise, including multi-language support; a proxy-server architecture to securely extend content management across geographically dispersed Web sites; and support for database distribution and replication capabilities through the Oracle Database. Based on J2EE, VIP Enterprise runs on the J2EE Certified Oracle9i Application Server. www.gaussvip.com

INTERLUCENT INTERNET SOLUTIONS INTRODUCES IBABBLER

3/26/2002

Interlucent Internet Solutions, Inc. announced the release of a developer tool that will reduce the implementation and conversion process of large websites into a content management system. Interlucent iBabbler is a Microsoft Windows application that spiders existing sites and translate them into new sites in Interlucent iCMS. Users create rules that help determine what types of pages the spider has found and what in that page is relevant content. Since a translation occurs a whole new look and feel can be implemented while preserving the content. This is a multi-use application so large sites can be translated in sections by different users.

www.interlucent.com

APPLIED SEMANTICS ANNOUNCES AUTO-CATEGORIZER 1.1 & TAXONOMY TOOL

3/26/2002

Applied Semantics, Inc. announced it has launched Auto-Categorizer 1.1, which features a taxonomy administration tool to enhance its content categorization application. This product update provides users the ability to import, create and edit sets of categories within a variety of industry-standard and user-defined taxonomies. Utilizing Applied Semantics' ontology-based CIRCA technology, the Taxonomy Administrator in Auto-Categorizer 1.1 maps concepts in a simple, direct manner without requiring the development of extensive training sets to define each of the categories. Applied Semantics' customers can now directly map their custom terminologies to its two-gigabyte ontology, a continuously maintained and updated knowledge base containing millions of concepts and relationship links through the Auto-Categorizer tool. Auto-Categorizer 1.1 with Taxonomy Administrator is available for immediate shipment.

www.appliedsemantics.com

TCP-IP XT NOW ENABLES REMOTE CONTROL OF ATOMIK 3.0

3/25/2002

Easypress Technologies and Techno Design announced support for Atomik 3.0 Developer Edition in the new version of Techno Design's remote control automation software, TCP-IP XT 1.4. TCP-IP XT is a platform, location, database and development environment independent scripting command set that supports over 320 commands for controlling the QuarkXPress user-interface functionality through a remote TCP/IP connection. The new version includes commands for controlling Atomik 3.0 Developer Edition, thus enabling automated QuarkXPress-to-XML conversion, remotely, over any TCP/IP connection. Atomik 3.0 Developer Edition is available direct from Easypress Technologies and Atomik resellers and system integrators worldwide. The suggested retail price for Atomik 3.0 Developer Edition is £14,995, \$19,995 or 24,292. Multi-user pricing is also available. Easypress Technologies is still adding overseas resellers and system integrators for Atomik. www.techno-design.com, www.easypress.com

STELLENT RELEASES ENHANCED OUTSIDE IN XML TECHNOLOGY

3/25/2002

Stellent, Inc. announced the release of Outside In XML Export 2.2. Outside In XML Export is designed to enable application developers to integrate XML conversion functionality into their solutions. Version 2.2 includes Stellent's new XML schema, SearchML, which is designed for information extraction in indexing, knowledge management and content management applications, and is provided with either a Java or C API. In addition to SearchML, XML Export includes FlexionDoc, an XML schema that allows for richer document conversion in more complex applications such as portals and content management systems. Outside In XML Export converts a document's content, structure and property information to one of these two schema. Outside In XML Export is a server-side technology with both a Java and component application programming interface (CAPI) that ensures ease of integration and implementation. www.stellent.com

ILANGUAGE FORMS PARTNERSHIP WITH ONLINE TRAINING FIRM

3/22/2002

iLanguage has identified multilingual eLearning as a major growth area for translation and localization companies, and has formed a partnership with IDON EAST of Newfoundland to offer eLearning solutions in multiple languages across Canada. In addition, iLanguage announced that it is planning to form similar partnerships with eLearning firms in the United States. The two companies will promote each other's services with existing clients, as well as new clients. This alliance will give both IDON EAST and iLanguage the opportunity to pursue clients that need both services, as a joint venture. www.idoneast.com, www.ilanguage.com

TRADOS ANNOUNCES RELEASE OF MULTITERM iX

3/21/2002

TRADOS Incorporated announced the official release of MultiTerm iX, the next version of their corporate language solution. Built around a multi-tier Internet architecture, MultiTerm iX is a flexible and scalable terminology management system. It is a database solution designed to standardize terminology and distribute it throughout the enterprise over the Internet or intranet at the click of a button. www.trados.com

VIGNETTE TO INTEGRATE V6 WITH BEA WEBLOGIC PORTAL

3/20/2002

Vignette Corp. announced that it is deepening its global strategic alliance with BEA Systems Inc. and will deliver a set of portlets that integrate its recently announced Vignette V6 MultiSite Content Manager (VMCM) with BEA WebLogic Portal 4.0. By using the portlets, VMCM will support and integrate with the BEA WebLogic Portal and enable current joint customers to extend their current investments in BEA technology. Vignette V6 MultiSite Content Manager, an extension to Vignette V6 Content Suite, enables organizations to manage content on multiple sites and portals through a single application. The VMCM application will help organizations maximize current or future BEA infrastructure investments to provide functionality such as single sign-on, layout and rendering of portal pages. www.vignette.com

XMLCITIES ANNOUNCES SUPPORT FOR ORACLE9I RELEASE 2

3/20/2002

XMLCities, Inc. announced it is using Oracle XML DB, a feature of Oracle9i Database Release 2, for a new application, XMLCapture Suite. Designed to automate the mass conversion of documents and images into XML for document archiving and information management applications, XMLCapture Suite, in combination with Oracle9i Database Release 2, now provides an integrated XML conversion and database solution. Oracle XML DB makes Oracle9i Database Release 2 a repository for XML content that is easily accessible by FTP, WebDAV, and HTTP. Oracle XML DB supports foldering, versioning, and access control of XML content stored and managed in Oracle9i Database. New features that have been integrated into XMLCapture Suite include advanced table handling, including CALS table standards, HTML output and workflow management integration. www.xmlcities.com

DOCENT & INTERWOVEN PARTNER

3/19/2002

Docent Inc. and Interwoven Inc. announced an alliance to improve the way knowledge is created, managed, and delivered in large enterprises. Through the alliance, Docent and Interwoven customers will be able to converge their organization's enterprise content strategy with their eLearning strategy. By integrating Docent Enterprise eLearning applications with Interwoven TeamSite content management software, customers can form an enterprise-wide content management infrastructure for knowledge exchange and more easily leverage information residing in content repositories. The ability to easily leverage the organization's content assets into multiple learning experiences extends the value of that content. www.interwoven.com, www.docent.com

DOCUMENTUM & GROOVE TO DELIVER SECURE CONTENT MANAGEMENT

3/19/2002

Groove Networks Inc. and Documentum announced an alliance to deliver a content management solution that provides secure content creation and management capabilities for project teams whose members span company boundaries. Both companies will market and sell the combined offering. The joint solution consists of a connector between the Documentum ECM platform and Groove decentralized collaboration software, which allows project teams to access content stored within the Documentum content repository and pull it into a Groove shared space. Team members can collaboratively view, edit and update the content - either online or offline. The project team can then publish final content in the central, secure Documentum re-

pository, where it can be accessed by other team members or anyone with appropriate permissions. www.documentum.com

IPHASE UNVEILS ONE STEP 3.0

3/19/2002

iPhrase Technologies, Inc. announced the availability of iPhrase One Step 3.0, the latest version of iPhrase's product for accessing high-value enterprise content. iPhrase One Step 3.0 brings to market a search and navigation system that can be implemented in just a few days, and fully deployed within 4-6 weeks. One Step 3.0 also allows enterprises to customize the results of queries based on the users' entitlement status, allowing organizations to personalize content for different constituencies. One Step combines retrieval of unstructured Web content and live access to structured databases into a single interface. The spider ensures deep content coverage, handling complexities such as frame sets, JavaScript, forms and dynamic URLs. Rich workflow integration with content management systems and external knowledge tools, such as Interwoven's TeamSite and MetaTagger, allows for exposure of all content. New in One Step 3.0, results can be personalized based on the user's entitlement status, enabling a single search solution to meet the needs of various constituencies. One Step supports XML, SSL, CSS, Microsoft .NET/Web Services, SOAP, ODBC, ASP and Java. In One Step 3.0, operating system support has been extended to include Sun Solaris and FreeBSD, as well as Windows NT/2000 and Red Hat Linux. www.iphase.com

INFODATA ANNOUNCES DOCLUMINATE

3/19/2002

Infodata announced the development of DocLuminate, a search and retrieval highlighting tool that allows term hit highlighting on Adobe Acrobat documents returned by web-based full text searches of documents in a Documentum repository. When a user searches on a specific keyword in an Adobe Acrobat document while using Documentum, DocLuminate highlights the term in the document. DocLuminate is compatible with Documentum's RightSite client, and Adobe Acrobat 4.0 or later, Adobe Acrobat Business Tools 4.0.5 or later, or Acrobat Reader 4.0 or later. DocLuminate will be available to the public in April and can be purchased through Infodata. www.infodata.com

IMANAGE ADDS NEW GOVERNMENT SECTOR PARTNERS

3/19/2002

iManage, Inc. announced new partners in the government sector that will support the company's penetration in local, state and federal government departments. Computer & Hi-tech Management, Inc. (CHM), and Ciber, Inc. will use iManage WorkSite to meet the requirements of their public sector clients for centralized, secure collaborative content management. iManage provides the government sector departmental and enterprise solutions that meet regulatory requirements for content management, including compliance with U.S. Department of Defense (DoD) Directive 5015.2, the Government Paperwork and Elimination Act (GPEA), Health Insurance Portability and Accountability Act (HIPAA) and other standards. The company's partnership with CHM and Ciber will help iManage deliver solutions and services to help government organizations address initiatives such as the Freedom of Information Act, correspondence management, e-government, case management, knowledge management and branch management. www.imanage.com

INTERWOVEN INTRODUCES CONTENT SERVICES SPECIFICATION

3/18/2002

Interwoven, Inc. announced the publication of the Content Services specification, a Web Services specification for access to Enterprise Content Management functionality. This proposed standard by Interwoven has support from vendors with portal offerings, including BEA, Bowstreet, Sun Microsystems, Inc., SAP Portals, Sybase and TIBCO Software, and will be submitted this month for consideration by one of the international open standards bodies. The new standard encompasses all lifecycle elements of enterprise content management, including content transfer, metadata, workflow, workareas, locking, and many other functions. These functions are made available as Web Services in a vendor-neutral solution, promising to hide many of the differences between competing software platforms, programming languages, and hardware specifics. The Content Services specification is available now through the www.content-services.org site for anyone to review, comment on, and implement. A full reference solution is included in the Interwoven implementation. www.interwoven.com

IDC SAYS CONTENT MANAGEMENT & RETRIEVAL SOFTWARE MARKET TO OUTPACE OVERALL SOFTWARE GROWTH BY 2006

3/18/2002

According to IDC, the content management and retrieval software market continues to grow rapidly despite the slowdown in 2001, and content management and retrieval software spending is expected to outpace the software market as a whole for the next five years. This growth has attracted the attention of large software vendors who are increasing their attention to this market. IDC's forecast predicts, that in 2003, pent-up demand for the next generation of content technologies, such as multimedia and multiformat search and text mining will fuel increased spending. According to IDC, this growth will only be temporary, followed by a shakeout ending with a few large players dominating the marketplace with unified enterprise information system offerings. The products that are easily integrated with other enterprise applications will have the greatest potential for success. www.idc.com

ECOSYSTEMS ANNOUNCES RELEASE OF THE ENVIRONMENT 1.7

3/15/2002

Ecosystems announced the release of The Environment 1.7 the latest edition of its XML content platform. The platform integrates with content management systems, databases and file systems to provide a single source representation of an organization. The software, available immediately, offers new functionality to improve document-publishing capabilities and content collaboration across the enterprise. www.eco-online.com

INTERMEDIA PHD LAUNCHES IQ AUTHOR

3/13/2002

UK-based internet developer InterMedia PhD launched IQ Author, its easy to use website content management product. With one copy of IQ Author, anyone can easily update their own website content; create pages; move pages around the site; delete old pages. There's no specialist knowledge required - content creation happens in a familiar word processor-like environment. With IQ Author you can add and delete graphics, change font sizes and colours, and alter the design of the website in just a few seconds. IQ Author includes as standard many of the "high-end" features seen in more expensive products. A single IQ Author powered website

can be managed by a team of people all who have their own centrally controlled access privileges. IQ Author is available immediately and starts at just £5000. www.igauthor.com

OBTREE TECHNOLOGIES EXPANDS TO U.S. MARKET

3/13/2002

Obtree Technologies has announced its expansion into the United States market with the opening of its Aliso Viejo, Calif. offices and the launch of its Obtree C4 architecture and Obtree Portal Suite solutions. The company plans to work with OEM partners to market its Obtree C4 architecture and Obtree Portal Suite, combined with other enterprise technology products to medium and large enterprise organizations in the financial services, pharmaceutical, and government markets. The offerings enable IT directors and Webmasters to integrate all enterprise content and knowledge management-related applications into one system that delivers needed information in real-time. Companies can use Obtree C4 and Obtree Portal Suite to manage, edit, and access legacy customer data, dynamic Web page content, market research reports, Excel spreadsheets, marketing presentations, and customized applications - all from one browser-based system. www.obtree.com

VIGNETTE TO INTEGRATE WITH WEBSPHERE PORTAL INFRASTRUCTURE

3/13/2002

Vignette Corp. announced it has developed a series of portlets that provide companies running IBM WebSphere Portal with access to Vignette applications and Vignette managed content. Vignette has developed these portlets to enable its recently announced Vignette V6 MultiSite Content Manager (VMCM) to integrate with the IBM WebSphere Portal. Vignette V6 MultiSite Content Manager, an extension to Vignette V6 Content Suite, enables organizations to manage content on multiple sites and portals with a single application. By using the portlets to integrate Vignette and IBM software, customers can create, produce and deploy content within the IBM WebSphere Portal infrastructure. The Vignette integration portlets for IBM WebSphere Portal will be available by March 31, 2002. www.vignette.com

MICROSOFT & ALTOVA TO PROVIDE XML DEVELOPMENT TOOLS FOR SQL SERVER

3/13/2002

Microsoft Corp. and Altova Inc. released new features to support Microsoft SQL XML and relational data by creating XML views of relational data. The new XML Spy provides enhanced tools support for Microsoft developers and database administrators. XML Spy 4.3 integrated development environment (IDE)'s schema editor now includes full development support for XML schemas, including Microsoft SQLXML extensions, for use in mapping XML schemas to relational databases. New SQLXML developer tools are included with the XML Spy 4.3 Suite, which is available for download and purchase. The XML Spy 4.3 Suite is available as a free upgrade to existing XML Spy customers. www.xmlspy.com/download.html, www.microsoft.com/sql

DIVINE TO ACQUIRE DELANO

3/13/2002

divine, inc. and Delano Technology Corporation announced that divine has signed a definitive agreement to acquire Delano. divine will acquire all of the outstanding shares of Delano common stock for a fixed exchange rate in which divine will issue 1.1870 divine shares for each share of Delano common stock, which equates to approximately 51.55 million divine shares.

The deal, which will be structured as a plan of arrangement under Canadian law, will be subject to customary regulatory and court approvals, as well as the approval of Delano's shareholders. Shareholders holding approximately 13 percent of Delano's outstanding shares have already agreed to vote in favor of the transaction. Combined with divine's CIM offerings, the acquisition of Delano allows divine to provide fully integrated multi-channel e-marketing, inbound and outbound interaction management, self service and campaign management. www.divine.com

SDL & TRIDION ANNOUNCE PARTNERSHIP

3/13/2002

SDL International, announced a strategic partnership with Tridion. SDL International and Tridion will work together to provide solutions for content management and globalisation. Tridion DialogServer, enables enterprises to publish, manage and update their web content. It allows for the consistent, timely and cost-effective exchange of strategic content across multi-language websites and devices. SDLWebFlow extends Tridion DialogServer with corresponding capabilities for the management of multilingual content. The key benefit to customers combining the two solutions will be the ability to manage the operation of web sites and content in almost any international language. www.tridion.com, www.sdlintl.com

SEMIO ACQUIRES KALEPA NETWORKS

3/12/2002

Semio Corporation has acquired Kalepa Networks, Inc. Kalepa, based in Palo Alto, Calif. is a developer of distributed discovery software. Terms of the acquisition were not disclosed. This acquisition is intended to enhance Semio's position in the content categorization and information discovery markets through new technology that will enable a rapid shift of focus from centrally managed solutions to distributed departmental solutions, including those located at the edge of the enterprise. www.semio.com

MONDOSOFT ANNOUNCES ENHANCEMENTS TO ENTERPRISE SEARCH AND REPORTING

3/12/2002

Mondosoft announced enhancements to its site search and reporting solution, MondoSearch, to provide corporate customers with efficient and economical search and reporting across their intranets, extranets, and Web sites. With full support for Unicode, MondoSearch 4.4 provides search for 37 languages Worldwide including Chinese, Japanese, Korean and Thai, and offers user authentication functionality for secure content delivery based on user privileges. MondoSearch 4.4 has also been optimized for use with content management solutions. Especially designed to support Microsoft Corp's Content Management Server, the user authentication technology provides the ability for different end users to receive information based on their specific privileges and access levels. Title, department, geographic location, and customer or partner status could determine such user levels. MondoSearch's search solution is certified for Microsoft's Commerce Server and tested and integrated with Microsoft's Content Management server and Solutions for Internet Business product lines. www.mondosoft.com

EKTRON ANNOUNCES XML AUTHORIZING TOOL FOR BUSINESS USERS

3/12/2002

Ektron, Inc. unveiled a new Web-based authoring tool that enables business professionals to quickly and easily add XML tags to Web content in a WYSIWYG environment. At the core of Ek-

tron eWebEditPro+XML is Ektron eWebEditPro. Often referred to as "a word processor in the browser," Ektron eWebEditPro enables business professionals to author Web content (tables, text, images, hyperlinks, etc.) without seeing HTML. Web developers have added this component to applications for content management, HTML mail, and discussion forums. Ektron eWebEditPro+XML offers the complete functionality of eWebEditPro, plus the ability for users to easily and transparently add XML or custom tags. Developers who integrate Ektron eWebEditPro+XML into applications can choose from Ektron-provided samples, or they can create custom, organization-specific XML templates, buttons, and drop-down lists. Content contributors can then add XML tags by selecting a template and adding content. Ektron eWebEditPro+XML is immediately available. A 10-seat license is US\$599. www.ektron.com

BACKWEB ANNOUNCES PARTNERSHIP WITH OBLIX

3/12/2002

BackWeb Technologies announced it has signed a joint marketing partnership with Oblix, Inc. to bring together the BackWeb ProactivePortal Server offering with Oblix NetPoint web access and identity management solution. Oblix provides integrated access control and enterprise identity management. Oblix NetPoint's identity management system, COREid, provides an infrastructure on which other e-business applications can be built. BackWeb technology, which provides offline access to critical content, is complementary to Oblix NetPoint, which ensures that only authorized users are able to view published content. With BackWeb and Oblix's combined solutions, customers can ensure that up-to-date critical content is always available, both online and offline, through secure delivery. www.backweb.com, www.oblix.com

BLUE MARTINI SOFTWARE SHIPS BLUE MARTINI 5

3/12/2002

Blue Martini Software, Inc. announced the availability of Blue Martini 5. The new suite delivers new functionality, as well as support for the IBM platform through Blue Martini 5 - IBM Edition, which comes bundled with IBM WebSphere Application Server and IBM MQ Series on AIX. Blue Martini is delivering new B2B selling and content management capabilities in Blue Martini 5, including lead management, RFQ, return material authorizations and invoice presentment. The new content management features include dynamic publishing for the immediate publication of new content to a website, as well as enhanced remote user functionality such as remote workflow and content preview. Blue Martini 5 is comprised of four applications: Blue Martini Marketing, Blue Martini Commerce, Blue Martini Channels and Blue Martini Service. In addition to the four applications, Blue Martini Software offers content management and integrated analytics, which underpin the four applications. Blue Martini Marketing, Blue Martini Commerce, Blue Martini Channels and Blue Martini Service are currently available. Each application is available separately. Pricing starts at \$85,000 per CPU. www.bluemartini.com

CROWNPEAK TECHNOLOGY UNVEILS TEMPLATE MANAGEMENT SYSTEM

3/12/2002

CrownPeak Technology introduced a new high-level template management system that provides business users with the ability to change the look and feel of an entire Web site via a browser interface. The company also announced a browser-based development environment targeted at professional Web developers. This template creation and management interface enables business users to modify templates, preview the resulting changes, schedule those templates live, and redesign the site on the fly. Because output templates are managed as assets in the system, multiple new look-and- feel options can be created, previewed, run through workflow, offered for collaboration and scheduled live. More basic changes to an existing template

can be performed through a drag-and-drop interface. More complex actions, like adjusting navigation or setting a new layout, can be performed via a "helper" interface, or directly with the template's HTML code. Advantage CMS is interoperable with other systems and applications. Connection options available to developers include SOAP, XML, FTP, Web Services, SMTP and ODBC. www.crownpeak.com

MEDIABIN & INTER-SIGHT DELIVER SOLUTION FOR WEB & PRINT

3/12/2002

MediaBin Inc. and Inter-Sight Interactive Insight Network announced that they have developed an integrated, end-to-end solution that allows marketers to dynamically publish marketing communications materials to print or to the web. Inter-Sight's Inter-Q AutoPage software automates the detail-intensive production and proofing tasks required to create virtually any customized print, web or software catalog from a single set of publishing-ready product data and images. AutoPage caches data from corporate product databases and pulls images from MediaBin's central repository of approved, "gold-standard" marketing and brand assets and then outputs QuarkXPress, InDesign, Acrobat, HTML, or XML files. www.intersight.com, www.mediabin.com

STRATIFY LAUNCHES WEB SERVICES FOR UNSTRUCTURED DATA MANAGEMENT

3/11/2002

Stratify, Inc. announced Web services for unstructured data management. The Stratify support for Web services provides a standards-based interface to the core functionality of the Stratify Discovery System that organizes, classifies and presents large amounts of unstructured data within enterprises and government organizations. Stratify's new Web services interface are SOAP and WSDL compliant. With Stratify's support of SOAP and WSDL, companies can publish Stratify's Web services into a UDDI directory. Also, developers operating in either Microsoft's .NET or J2EE environments can now incorporate the Stratify Discovery System into their applications. www.stratify.com

DIVINE ANNOUNCES AVAILABILITY OF CONTENT SERVER 4.0 & PARTICIPANT SERVER 4.0

3/11/2002

divine, inc. announced the availability of divine Content Server 4.0, and Participant Server 4.0. Content Server 4.0 provides companies with a transparent growth path from point project to shared enterprise infrastructure for content management. Content Server also provides a transparent view of enterprise content to end-customers of a Web site or content-driven application, drawing data from existing IT systems and diverse formats and presenting it in a single branded interface. divine extends Content Server 4.0 by adding native support for WebLogic 6.1 and WebSphere 4.0. In addition, divine Content Server is available as an integrated part of the portal platforms, including the iPlanet Portal Server, BEA WebLogic Portal and the IBM WebSphere Portal Server. Also with the launch of Content Server 4.0, divine provides further support for XML Web services standards. divine Content Server 4.0 and divine Participant Server 4.0 are available immediately. Participant Server is available on Windows NT, Windows 2000 and Sun Solaris 2.8. Discounts for volume purchases and deployment licenses are available. www.divine.com

FILENET ADDS LDAP & J2EE SUPPORT TO PANAGON

3/6/2002

FileNET Corp. announced it has added LDAP and J2EE support for its Panagon Enterprise Content Management (ECM) solutions. The integration of these capabilities makes it easier for enterprise customers to drive eBusiness initiatives by delivering customer facing, Web-based content and process management applications with the safeguard of enhanced security. The addition of J2EE support enables organizations to bring high volumes of enterprise content into customer facing Web-based applications using Java development environments. FileNET will offer a complete set of Panagon J2EE capability. The first release is available immediately for Panagon Image Services, which can be used for the development and deployment of Web-based applications in Unix environments. Designed to meet the J2EE Connector Architecture (JCA) 1.0 specifications released by Sun Microsystems in July 2001, developers experienced with J2EE can easily use the Panagon resource adapter to access and leverage existing enterprise content with no additional training. The Panagon J2EE capability has no specific client-side operating system or browser requirements. www.FileNET.com

LIGHTSPEED INTERACTIVE & BARBADOSOFT BV PARTNER

3/6/2002

Lightspeed Interactive, Inc. and Barbadosoft BV announced that they have entered into a strategic alliance to promote one another's products in both the United States and European markets. Lightspeed believes that the co-marketing and co-selling arrangements with Barbadosoft will add significant value to its product offering because CorteXML eliminates the risks inherent in building and maintaining XML-based environments, allowing users to take full advantage of the benefits of XML. www.lspeed.com

ATOMZ ADDS XML DATA IMPORT & EXPORT

3/6/2002

Atomz Corporation announced support for XML data import and export. Now the data that Atomz Publish maintains for its customers can be exported to other applications in the enterprise, and data from other data repositories can be easily uploaded and imported into the Atomz Publish Web content management system. Atomz already provides an XML content repository with Atomz Publish. The feature enables users to define data schemas for both importing and exporting XML data directly to Publish templates. Administrator users employ XML syntax to define the relationship between the Atomz Publish data and the external data source in order to create data schemas. Complete with user permissions, multi-step tasks and email notification, workflow is easily configured and managed by the administrator, who can also define tasks to incorporate exporting and importing data. Uploading templates and data schemas can be done from the Administrator's desktop. Data fields can be re-ordered upon either data export or import. This enhancement also allows administrators to convert .html files to .xml files on export and conversely convert .xml files to .html files on import.

www.atomz.com

IMANAGE UNVEILS COLLABORATIVE CONTENT MANAGEMENT SUITE

3/6/2002

iManage, Inc. introduced iManage WorkSite MP (Multi-Platform). WorkSite MP is an application suite to combine document management, collaboration, workflow, and portal access in a single integrated Java-based solution that operates across Windows NT, Linux, and Solaris platforms. iManage WorkSite MP streamlines organizational administration by enabling companies using mixed business environments to utilize a single, integrated solution. WorkSite MP is designed

for a wide variety of applications across many industries, including supplier relationship management, contract management, asset and wealth management, wholesale banking and regulatory compliance. The iManage WorkSite MP application suite includes four components: WorkDocs MP, WorkTeam MP, WorkPortal MP and WorkRoute MP. iManage WorkSite MP is available immediately through iManage and its partners. The introductory price of WorkSite MP starts at \$70,000 for 100 users. www.imanage.com

IPEDO INTRODUCES XML DATABASE 3.0

3/5/2002

Ipedo, Inc. introduced the newest version of its Ipedo XML Database 3.0. The Ipedo XML Database's unique fusion of XML query, management and transformation capabilities gives enterprises an XML infrastructure available today for distributed, dynamic information management. The Ipedo XML Database 3.0 introduces fully W3C-compliant XML Query implementation, new XML schema evolution capabilities, unique XML document versioning, improved SOAP support and further performance and scalability improvements. With these enhancements, businesses now can have the dynamic and reusable XML content infrastructure they need to support .NET and other Web Services applications, as well as a variety of enterprise portal, mobile, supply chain, collaborative and other XML content-intensive and personalized applications. The Ipedo XML Database 3.0 is available now for Windows 2000, Windows NT, Sun Solaris and Red Hat Linux. Pricing is on a per-CPU basis. www.ipedo.com

GAUSS & ARTESIA PARTNER

3/5/2002

Gauss announced that the company has signed a Global Technology Partnership Agreement with Artesia Technologies, a market leader in Digital Asset Management. Under this agreement, the companies will work closely together to offer integrated, bundled solutions of their respective software products - VIP Enterprise 8 and TEAMS. The Gauss-Artesia partnership provides Global 1000 organizations with a unified solution spanning their diverse enterprise content repositories, allowing content contributors from a wide variety of areas to create, contribute, manage and dynamically repurpose all types of enterprise assets. The goals of the Gauss/Artesia Technology Alliance include product bundling and integration efforts, co-marketing campaigns and co-sales activities, initially in North America, and then in Europe. www.artesia.com, www.gaussvip.com

ACCUMEDIA ANNOUNCES ACCUMEDIA CONVERGENCE PLATFORM 2.5

3/5/2002

Accumedia announced the release of version 2.5 of its Accumedia Convergence Platform (ACP) for Media, a web services content management and delivery solution specifically developed for the media industry. This latest version of Accumedia's content management software provides print editors with the capability to easily schedule personalized content simultaneously for multiple publications or brands using simple, non-technical user interfaces. The new version's new features include: WYSIWYG Content Editing, Graphical Content Scheduling, Segmentation-based Content Delivery, In-context Preview, On-Line Subscriptions, and E-mail Marketing. www.accumedia.com

VIGNETTE ADDS SUPPORT FOR WEB SERVICES

3/5/2002

Vignette Corp. announced that it has added support for Web services to Vignette V6, the company's Web content management solution. Vignette V6 enables content management processes to be exposed as Web services by offering support for the XML Schema, WSDL and SOAP standards in new Web services and XML adapter libraries. Vignette V6 also offers support for major industry-standard platforms, including the J2EE and Microsoft .NET standards, in addition to the IBM WebSphere, BEA WebLogic and Sun iPlanet application servers. Vignette's Web services and XML adapter libraries are immediately available and are included with Vignette V6 software licenses in the Vignette V6 Content Suite Adapter. Existing customers can download the adapter libraries via Vignette Express, the company's electronic software delivery service. <http://global.vignette.com/>

CLEARFOREST ANNOUNCES CLARTAGS 4.0

3/5/2002

ClearForest announced ClearTags 4.0, a comprehensive auto-tagging platform. The new version, which includes semantic, statistical and structural tagging, expands tagging output and the understanding and value of unstructured content. A new user control panel allows the definition of different tagging schemes for any type of document stream, monitoring the entire tagging process. ClearTags allows publishers, content providers and other content-intensive businesses, to precisely identify and automatically tag multiple relevant entities, facts and events buried within large textual repositories. The process produces richly-tagged XML files. The output of ClearTags can be used to create new products and re-package content for various downstream applications and delivery methods, or for further analysis. ClearTags accepts input in a variety of formats, including PDF, MS Office, HTML, and XML, and automatically enriches each document with an extensive set of relevant meta-tags. The tags are based on three main technologies: Semantic/Linguistic Information Extraction, Statistical Categorization, and Topological Analysis. ClearTags is also used to generate a ClearForest knowledge base, to be used with ClearResearch, ClearForest's enterprise research application, or third-party Web applications. www.clearforest.com

DOCUMENTUM DELIVERS DIGITAL ASSET MANAGEMENT

3/4/2002

Documentum announced the immediate availability of two new product offerings, Documentum Media Services and the Digital Asset Management Edition. Both offerings are based on the Bulldog Group technology that Documentum acquired in December 2001. Documentum Media Services, which is integrated with the Documentum ECM platform, provides intelligent handling of rich media assets, such as media-specific indexing, browsing and search capabilities, easy-to-use editing, automated file transformation features, and the ability to fully integrate and manage these digital assets with other types of content. The Media Services product is targeted at the mainstream requirements of the Global 2000 and provides a cost-effective solution for managing rich media. The Digital Asset Management Edition offers advanced rich media management capabilities for complex, broadcast-quality digital and physical assets, such as high-resolution video and is targeted at high-end rich media requirements. www.documentum.com

INTERWOVEN & NEC TEAM

3/4/2002

Interwoven, Inc. announced a strategic alliance with NEC Corporation whereby NEC will re-sell the Interwoven Content Infrastructure product suite as the core enterprise content management functionality in its iBestSolutions. NEC's Internet solutions framework, iBestSolutions, consists of the complete line of NEC's software and services offerings. Under the new reseller agreement, NEC will adopt Interwoven's TeamSite, and the Interwoven Content Infrastructure suite, as the exclusive enterprise content management system underpinning its own customer relationship management (CRM), enterprise information portal (EIP), and eCommerce (EC) solutions. NEC also plans to develop integrated solutions that incorporate Content Infrastructure into its existing product offerings, establish an enterprise content management training and support program and initiate co-sales and co-marketing activities. The focus throughout 2002 will continue to be on solutions for the manufacturing sector, with a view to expanding to additional industry areas in the future. www.nec.com, www.interwoven.com

THE GILBANE CONFERENCE ON CONTENT MANAGEMENT @ SEYBOLD

September 11-12, San Francisco, Moscone Center

Join us for our newest conference, where all the major content management players, as well as consultants, analysts, and experienced IT and project managers will share content management strategies, learn about new content management technologies, and debate approaches.

Being co-located with the Seybold San Francisco 2002 means you can explore a whole range of content technologies in one place, including XML, DRM, DAM, PDF, *etc.*

Conference description

Content management has evolved from an activity mostly associated with managing Web page content into a mainstream IT discipline focused on managing all enterprise content. Enterprise content includes *unstructured* data, such as email, Office documents and rich media, *structured* data from relational databases; and XML data, which is being now used to encode and share all types of content. The most difficult aspects of implementing a successful content management system today are *integrating* it with other enterprise applications and content, *organizing* the content so it can be located, and *communicating* it to the appropriate audiences. Our conference is focused on these three areas.

Why?

Managers of content management systems, corporate content strategies, and IT infrastructure strategies all need to understand what content management technologies can do for their business, and how to evaluate the hundreds of new products and technologies available managing and integrating enterprise content. Our conference will provide you with the fundamental knowledge you need to build a successful strategy and implementation.

More info & registration:

<http://www.gilbane.com/seyboldsf02.html>

<http://www.seyboldseminars.com/sf2002/>

BACK ISSUES

Issues from 1993 thru 1998 are \$15 if in print. More recent issues are available in PDF for various prices and *may* be available in print form for \$30. See www.gilbane.com or call for more information.

Volume 10 — 2002

- Number 2 **The Many [Inter]Faces of Content Management Systems**, *Rita Warren*
Number 1 **What is an Information Model, and Why Do You Need One?**, *JoAnn Hackos*

Volume 9 — 2001

- Number 10 **An Alternative Model for Personal Information Management**,
Girish Altekar
Number 9 **Who Should Own Your Content Management System?**, *Bob Boiko*
Number 8 **Understanding Web Services**, *Sebastian Holst*
Number 7 **Editorial Interfaces & Enterprise-enabled Content**, *Bill Trippe & David R. Guenette*
Number 6 **Why Content & XML Integration Technologies are Fundamental**,
Frank Gilbane
Number 5 **The Application Server Cometh, II**,
Bill Trippe
Number 4 **Open Source Content Management Systems: A Parallel Universe?**,
Sebastian Holst
Number 3 **Privilege Management & Rights Management for Corporate Portals**, *David R. Guenette, Larry Gussin, and Bill Trippe*
Number 2 **Choosing an Architecture for Wireless Content Delivery**,
Girish Altekar, Regan Coleman
Number 1 **XHTML: What You Should do About it, and When**,
Sebastian Holst, David R. Guenette

Volume 8 — 2000

- Number 10 **XML: The State of the Union** *Bill Trippe, David R. Guenette*
Number 9 **E-books: Technology for Enterprise Content Applications?**
Bill Trippe, David R. Guenette
Number 8 **What is Content Management?** *Frank Gilbane*
Number 7 **Syndication, Actionable Content and the Supply Chain**,
Bill Trippe, David R. Guenette
Number 6 **Digital Rights Management: It's Time to Pay Attention**, *David R. Guenette*
Number 5 **E-catalogs: Strategic Issues for Suppliers**, *Frank Gilbane*
Number 4 **Content Management: Application vs. Application Server Solutions**,
Bill Trippe
Number 3 **XML on the Front End: Connecting People & Processes in B2B E-commerce Environments**, *Mary Laplante*
Number 2 **ASPs, Content, & Code**, *Frank Gilbane*
Number 1 **XML, EDI, Content, & Commerce**, *Bill Trippe*

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

Website Content Management. May 27-28, Post-Conference Workshops, May 29, Holiday Inn Victoria, London UK. An IQPC event. "Discover and discuss how successful website content management can increase profits, brand awareness and improve information distribution." ***Gilbane Report subscribers receive £100 discount!*** Enter the Gilbane Report member's section for instructions on registering for the discount. www.iqpc.co.uk/GB-1801/gilbane

Content Management: Strategies for Single Sourcing. June 3 and 4, Hyatt Regency, 5 Embarcadero Center, San Francisco, CA. Presented by JoAnn Hackos and The Center for Information-Development Management. If you're trying to manage corporate data, reuse your content, and deliver information through many media, join your peers at this event. Successful content management results in reduced costs and ROIs. Learn how you can realize both for your company. Keynote speakers: Dr. JoAnn Hackos, Dr. Jonathan Price, and Frank Gilbane. www.cm-strategies.com

Establishing the Business Benefits of Intranet Content Management. June 26 and 27, Hilton London Olympia – London UK. At this event, Ark Group will address all of the challenges you are currently facing in terms of quantifying the ROI of Intranet content management. ***Special 10% discount for Gilbane Report subscribers!*** To register or to download a pdf brochure, visit www.ark-group.com/fg-offer.html.

The Gilbane Conference on Content Management @Seybold. September 11-12, 2002, Moscone Center, San Francisco, CA. Our newest conference will be co-located with Seybold San Francisco 2002 so that IT and project managers from all industries now have a one-stop event where they can learn how to build content management strategies, and create implementation plans that integrate with other enterprise applications in order to meet a wide variety of business needs. www.gilbane.com/seboldsf02.html (the official site will be up soon at www.seyboldseminars.com/sf2002/)

© 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