Best Practices in Enterprise Knowledge Management

Andy Moore ............................. 2  
Passing the Cringe Test: Has KM Made It to Prime Time?  
Once upon a time, saying the words “knowledge management” was the fastest way to get thrown out of a meeting. And not only would you be ejected, you would never be invited back. . . .

Paul Sonderegger, Endeca ................ 4  
People Judge Relevance. Machines Calculate Evidence.  
Does the man make the machine? Or the machine make the man? A pilot can’t fly without a plane. A doctor can’t diagnose without a centrifuge. . . .

Daryl Orts, Noetix .......................... 6  
How to Correctly Design and Implement a Dashboard  
A dashboard is a vital tool for monitoring the daily health of your organization. From a single interface, decision makers have access to key performance indicators (KPIs). . . .

Hadley Reynolds and Silvija Seres, FAST ........ 8  
Building the Search Center of Excellence  
Search is strategic; however, the strategic potential of search is not captured by the act of acquiring a powerful search platform alone. . . .

InQuira .................................................. 10  
Implementing KM: Practitioners Share Best Practices  
Most customer service organizations today will admit that effective knowledge transfer is the most crucial element to resolving customer problems. . . .

Brandon Lackey, BEA, and Michael Behounek, Emerja ............ 12  
Mistakes to Avoid and Principles for Success  
Many organizations today still struggle to get value from their knowledge management (KM) efforts. Even companies that have been benchmarked as the “best practice” can easily fail. . . .

Jeff Dirks, SchemaLogic ..................... 14  
Knowledge is Power: Best Practices in Semantics Management  
Knowledge is power. The meaning of the phrase is even more applicable today as companies around the world are dealing with an overwhelming amount of data. . . .

Anand Chopra, KANA .......................... 16  
Answering Customers’ Questions the Intelligent Way  
Enterprises face a difficult challenge when simultaneously improving the quality of service and reducing costs. More products and rapid change substantially increase the amounts of information. . . .

Gordon Taylor, TOWER Software ................ 17  
Knowledge at the Point of Decision  
A small, two-legged robot stands atop a glass-topped coffee table. On its two dimensional world, it has to contend with a potted plant, an old TV Guide and several coffee cups. . . .

Laurent Simoneau, Coveo ..................... 18  
Enterprise Search: The Foundation for Risk Management  
If you can’t find it, you can’t manage it. Unfortunately, many executives have discovered this too late in the game—when they face legal troubles or a corporate scandal. . . .

Jon Parsons, XyEnterprise ..................... 19  
Realizing Measurable ROI with Multi-Language Content Management  
What is content management worth? That can be a philosophical question. If you ask a content author, you’ll hear about the way content creation and review is made easier. . . .

Dan Dube, DocZone.com ....................... 20  
Hosted XML Content Management: Is It Right for You?  
Publishing organizations have long recognized the value of migrating content to XML to attain the benefits of content reuse, reduced localization costs and single-source publishing. . . .

Gregory F. Roberts, Lockheed Martin ............. 21  
Who is that “he”?—Using Pronouns and Anaphors in Text Extraction  
Text extraction is a powerful tool to find and categorize elements in unstructured documents. These elements, or entities, are connected together to form the relationships, facts and events. . . .

Andrew Cohen, KNOVA Software ............. 22  
Beyond First Call Resolution  
Knowledge management (KM) initiatives are one way of improving technical support organizations, driving support margins through efficiency and increased customer loyalty. . . .

Chip Gettinger, Astoria Software ............. 23  
The Hidden Costs of Product Information Publishing  
Manufacturing enterprises have placed a sustained focus on information management solutions to support and augment the design, development and production process. . . .
Passing the Cringe Test

Has Knowledge Management Made It to Prime Time?

By Andy Moore, Editorial Director, KMWorld Specialty Publishing Group

Once upon a time, saying the words “knowledge management” was the fastest way to get thrown out of a meeting. And not only would you be ejected, you would never be invited back.

That was then, but this is now. I asked a panel of experts not only whether KM was finally ready to pass the “cringe test,” but what it would take to propel KM to even greater prominence as an enterprise set of solutions. And I was surprised—and pleased—at the reactions.

“One of the big failures of knowledge management, and the reason it had the reputation it had, is that it tries to solve problems at too much of a macro level,” said Brandon Lackey. Brandon is the global solutions director for BEA Systems. “They tried to build a database of everyone’s knowledge, and it can’t be done. The successful systems may go in under the basis of knowledge management, but they address very specific business problems.”

So can you “talk about KM” during these meetings? “You can, in pockets,” noted Paul Sonderegger, the principal strategist and “evangelist” at Endeca. “There are companies that are more dependent than others on the movement of information in order to affect the bottom line. Take a professional service organization—a consulting company, for instance. Its whole ability to leverage the work their consultants do depends on their ability to take information others have created and apply it in another context. In that kind of environment, yes...you can absolutely talk about knowledge management.”

Tim Shetler, vice president, marketing at InQuira, thought his company knew the difficulty in the terminology: “We deliberately named our product ‘Information Manager’ to avoid any issues with the term ‘knowledge management.’ We discovered they don’t exist! We thought there would be some tarnish around the term. But every RFP I look at uses the phrase ‘knowledge management.’ The end users don’t have that baggage.”

It’s probably a matter of whom you ask, and what type of organization they work for. “The term ‘knowledge management’ doesn’t make sense to a guy pouring concrete,” insisted Sally Hicks, marketing manager at Noetix. “Our biggest dashboard customer is Florida Rock. Their name says it all—they move rocks. It’s not sexy. But they allow their customers to log in through a secure dashboard and get a few basic pieces of data, and it’s saving hundreds of man-hours per week. Their customers don’t know the terminology of ‘knowledge management.’ They just want a copy of an invoice.”

“We are able to now talk about knowledge management without the audience rolling their eyes,” said Sivija Seres, FAST’s VP of strategic business development, “but we have learned a lot. People have been told one thing at a marketing level, but it’s difficult to achieve those things when it scales greatly and gets more complex, and more features creep up into the original specification,” said Sivija. “Recently, the technology—not the core, but the functionality—and our understanding of how to set things up from the beginning has reached a level that we can be really proud of.”

Unmet expectations is a common theme in technology deployments, of course. But maybe it’s worse in KM, because the goals are often so ambiguous. “The problem with asking users to describe what they want is like the old saw about Henry Ford,” remarked Paul Sonderegger. “If Ford had asked people what they wanted, they would have said ‘faster horses.’ Luckily, there is a rising generation of managers who have a greater baseline of familiarity with technology than the departing generation. They’re not technologists, but they have new ideas about what they want from their infrastructure.”

Lowell Anderson, VP marketing at SchemaLogic, agreed with that assessment. “There’s definitely an increased awareness of the capabilities that a focused KM strategy can bring to an organization. As we become familiar with collaboration techniques, people are becoming more comfortable with moving away from fixed standards as a way to govern knowledge. Open collaboration effectively becomes the standard,” said Lowell.

The KM Formula

If KM is now accepted among polte company, what’s holding up its widespread adoption? “We always come back to the same fundamental thing: people can’t get the information they need to do their jobs,” said Sally Hicks. “Even if IT is helping get information to people, they don’t get it quickly enough, or it changes. Sometimes the users don’t even know what it is they want; they just know they need something in order to make a decision.”

The theme of “information+decision=knowledge management” is practically universal. “What’s the connection between information and action?” asked Paul, entirely rhetorically. “It’s the assessment of the problem and the appropriateness of available solutions. Before you can do anything logistically—like shipping something from one place to another to fix a problem—someone has to assess the problem, and become satisfied they have identified the right solution. And that’s ALL about decision-making. It’s the information being brought to the person that allows us to talk about the application of knowledge.”

Of course, as is often the case, it’s easier said than done. “As customers spend significant amounts of money deploying information-access networks, they are realizing they aren’t working as well as they...
possibly could,” said Lowell. “They are connecting different repositories and content silos, but they don’t have a common understanding of the terminologies that describe the information residing in those content silos. There are also complex interrelationships among those terms they have no way to model.”

Sally Hicks agreed: “The terminology issue is huge. If you’re working with a database, and there isn’t a term that exactly matches ‘table 1’ in the database, there is absolutely no way you’re going to get what you need. In many cases, the user isn’t even sure he HAS the information, much less be able to find it using strict database terms.”

And it doesn’t appear that the average technology tools provide much help; a more integrated solution is required. “A lot of customers feel that ‘search’ by itself is not as big a problem as content or knowledge,” said InQuira’s Tim Shetler. “Most of our customers now come to us looking for both. It’s a much broader solution area.”

“A useful lens to look at this through is the long tail,” said Paul. “Within an enterprise, there will be a number of requests for information that everyone asks all the time—total blockbusters. Then there’s the body of requests that are kind of in the middle, relatively common but not blockbusters. But then you’ve got an almost infinite number of requests that occur seldomly... maybe only once. So any kind of ‘master schema’ will fail.”

Getting in the Door

Perhaps because of the complexity; perhaps because of the ambiguity of the benefit, KM is a tough sell. “Selling knowledge management systems is harder than selling systems that help people make money,” said Silvija. “You have to provide a very strong ROI argument. You have to help customers understand how it will either make them more effective, or save money,” she said.

“The difficulty begins when you want to roll it out. You have to convince the IT person at every division or office. For them this is just a headache; they have to replace something that’s already in place and working with something that will take time to get right. In order to grow a system to a very large scale with very advanced features, you need to have a nicely staged process, where they first get something that works really well at a basic level, then add new features and grow the system in a controlled way.”

“What it works right,” added Brandon Lackey, “IT recognizes that the business problem has to be solved by the line-of-business stakeholders. IT should coordinate, and see themselves as the enablers. So much of knowledge management is people and process; if the business says to IT ‘Make me a repository,’ IT should push back and say ‘No, this is so much more than that. You need to make sure the processes are open and the right people are in place.’” All agreed that Brandon’s view is correct, but utopian. “Business wants a quick fix, and that’s been the failure of knowledge management: there’s no one.”

“IT will end up owning and running the system, so it’s OK for them to be in charge,” added Sally. “But they need to have the input and buy-in from business, and listen. And the end users also have to take the time to work with IT and ask for exactly what they want up front.”

“Every RFP uses the phrase ‘knowledge management.’ The end users don’t have that baggage.”

“The more ambitious of our customers know they have to view it as cross-functional and cross-divisional. They see it as a platform, and on top of that they keep growing projects,” explained Silvija. “They are starting centers of excellence in greater numbers. But in order for them to succeed, you need support from quite a high level.”

What’s Next?

Tim Shetler detects a “mood shift” among KM system users and purchasers: “The resurgence of KM is driven by the distribution of workers, and work, both geographically and over time. For companies like that, knowledge management is something you have to do just to run the business. But there’s another change: people are looking for ways to affect their top line. Sure, people are still looking for ways to cut costs, and get ROIs from cost centers, such as their call-center operations. But people are tired of doing that; they’re looking for ways to raise the top line—‘Give me a 2% increase in marketshare.’”

He continued: “An enterprise knowledge-sharing initiative has a lot of hard-to-quantify, soft-dollar benefits. That’s hard to justify. But if you can demonstrate a hard-dollar return, it’s much easier to spread when the case is proven.”

The most interesting turn in our conversations revolved around the opportuni-
Is Relevance Irrelevant?

The familiar framing of the information-access problem is how to help people find the right information at the right time. But we need some measure of “rightness” before we can characterize the tools we need. “Relevance” is the technical term for the central measure of “rightness” in information retrieval. Unfortunately, relevance defies logical rules. You know this if a search engine ever returned to you a document that was “72% relevant.” This is like a flight computer saying the plane is 72% on course.

Relevance is subjective because it’s relative to the person’s goal. For example, if I tell you that the movie tonight is at 7:30 PM, is that relevant? If we’re going to the movies, then yes. If not, no. How can a piece of information be both relevant and irrelevant? It depends on what the user is trying to accomplish. Relevance is in the eye of the beholder.

Just because relevance is subjective doesn’t mean we must give up hope in finding it. On the contrary, it must be the central focus of information access technology. But it does reveal why we shouldn’t ask the machine to determine relevance for us. It can’t. Fortunately, people are experts at judging relevance—and with help from machines, we can do better than without. The central function of information access technology is to produce evidence that informs human judgment. We summarize this as:

◆ People judge relevance
◆ Machines calculate evidence

Context Informs Relevance

So what can a technology do that would help people better determine relevance? First, let’s look at how context helps people determine meaning. Take this simple exchange:

Host: “Would you like some dessert?”
Guest: “It would fill me up.”

How does the host figure out if the guest wants dessert or not? If the host knows that his guest doesn’t like to sleep on a full stomach, then the answer is no. If the host knows that his guest likes to eat until satisfied, the answer is yes. This context lives outside the literal dialogue in the mind of the host. It completes the proper meaning of the guest’s statements. But how does the guest know the host will make the proper interpretation? He’s actually relying on the host to have this context in mind and, therefore, interpret his statements accurately.

This is how human communication works. It’s based on the resolution of ambiguities through best efforts and educated guesses. But both parties have to be working with the richest context-processing machine available—the human mind. Both guest and host use observations about the current environment, knowledge of the audience and personal experience to attempt to communicate effectively. The vast majority of the time it works. And the application of these contextual facts, taken from the environment and the head, is the key.
If we are to make humans the determinant of relevance in an information access application, it is context that helps them judge information. But how do we get from content to context?

**Relationships in the Content Inform Context in the Head**

In the man/machine exchange of information access, only one party is reasoning with the flexibility and inventiveness of a human being. To make the machine play a helpful role in the dialogue, there are two basic approaches:

1. Mimic human judgment with software rules
2. Use software rules to inform human judgment

The first works well in constrained, static user scenarios. For example, a digital camera sensing low light will leave the “shutter” open longer, just as a photographer would. But there’s no setting for capturing, say, a child’s joy of accomplishment. And if there were such a setting, would you trust it on the day of your child’s big game? Search engines attempting to calculate relevance are over-reaching just as much. These algorithms are really just counting—whether it’s the number of times the query term appears in the target documents, the number of links using that term which point to a given Web page, or something based on probabilistic or other statistical calculations. In each case, the raw calculations fail to capture what the indexed information is about. Worse, the mechanism is a black box. When the user gets results that appear irrelevant to him, it’s hard to tell why his best-attempt query failed.

The second approach works well in dynamic, unpredictable scenarios. Think of a doctor in an emergency room. A patient arrives complaining of weakness, shortness of breath and numbness in his extremities. A blood test shows low oxygen saturation and a chest x-ray shows fluid in the lungs. Neither of the machines that produce this diagnosis can do the conclusion that the patient was poisoned. But the doctor can. The tools provide the doctor with facts that better inform her judgment. But it’s the doctor who puts the whole picture together. In information-access technologies, databases and BI tools are much closer to this approach. They store data in specific structures, allowing them to maintain the factual relationships inside. The key is to point the machine at the smallest indivisible unit of information and then give the user the tools to see those pieces as a whole picture.

That unit is the facet—an explicit characteristic of a thing. Facets are everywhere—metadata on documents, user-contributed tags and fields in a database record are all facets. And in a strange twist, the full-text index of a document can be thought of as just a dynamic facet of that file.

But don’t relational databases work this way today? Not quite. People easily conceive of dimensions that cut across a “rectangular” database. Take a customer database; it’s straightforward to find out which customers bought a particular product in the past year. But it’s much more difficult to find out which combinations of products were bought in the last year by customers in a particular region—a sensible question, based on dimensions that cut across the original database structure. That’s what BI is for. Unfortunately, BI can only support the questions the developers knew ahead of time would be asked. But people are much more flexible than that. They come up with unforeseen questions and unpredictable ways to express otherwise mundane requests. And, they want to take into account information in different forms from records to documents. That’s what search is for. But, of course, traditional search engines don’t maintain relationships from the systems they index. And we’re back where we started.

**A New Architecture Calculates Relationships Among Relationships**

Just as modern planes, medical technologies and cameras give pilots, doctors and photographers more flexibility and control, a modern information-access platform must do the same for employees, customers and partners. Such a machine must:

- **Accept messy, real-world enterprise information.** The platform must be able to index enterprise data and content that comes in all its different formats, sizes and quality.
- **Preserve all the relationships in the original systems.** The relationships in the databases, enterprise apps and documents are pre-existing investments the platform must exploit.
- **Calculate all the dimensional relationships.** The facets in each record and document are the basis for connections across otherwise completely different assets.
- **Guide users through constantly shifting contexts.** Each time the user takes a step through the app—whether a keyword query, a navigation selection, a circled region on a map—the platform must show him the results and all the possible, but only the valid next steps.

These requirements call for an architectural capability called *adaptivity*. Adaptivity is the dynamic calculation of relationships among relationships in the current results set based on the possibilities in the data, the user’s actions and any business rules. Information access applications built on such a platform show a user all the results to a query plus all the information about those results, creating greater context for the user’s judgments. For example, a product engineer looking for titanium bolts for a lightweight chassis design searches for “titanium bolts.” The results include all the titanium bolts, of course, as well as all the dimensions by which this list of bolts could be refined—length, weight, finish, thread pitch and so on—plus the quality-assurance reports including data on titanium bolts, with the appropriate refinement dimensions.

And as the engineer selects refinements or searches within the results, or both, all the information on the screen dynamically updates, driven entirely on the possibilities that exist in the data. The result is the richest possible presentation of the evidence, guiding the engineer to his goal.

Machines should not pretend to do things they can’t. And they can’t judge relevance for a particular user with a specific goal. Instead, the machine should calculate the relationships among relationships in the data and content, fueling the context for decisions, making experts of us all.

Endeca, headquartered in Cambridge, MA is a next-generation information access company uniting the ease of search with the analytical power of business intelligence. The Endeca Information Access Platform combines patented intellectual property, breakthrough science and a deep focus on user experience to help people find, analyze and understand information in ways never before possible. Leading global organizations like ABN AMRO, Boeing and Cox Newspapers rely on Endeca to increase revenue, reduce costs and streamline operations through better information access.
How to Correctly Design and Implement a Dashboard

A Step-by-Step Guide

By Daryl Orts, VP of Advanced Solutions, Noetix Corp.

A dashboard is a vital tool for monitoring the daily health of your organization. From a single interface, decision makers have access to key performance indicators (KPIs)—actionable information that can be used to effectively guide and track business performance.

At a high level, it may seem relatively easy to build a dashboard. Companies that feel they have a good handle on which performance indicators are of strategic importance to the organization may think collecting, summarizing and consolidating the supporting data shouldn’t be that difficult. However, such oversimplification can lead to a failed project before it ever gets off the ground.

The successful implementation of a dashboard is complex and requires a step-by-step process: a methodology that considers all aspects of the project life cycle. This series of tasks—plan, design, build and deploy—will be similar, regardless of the technology or vendor chosen. When comparing proposals from multiple vendors or the cost of a “do-it-yourself” project, it is important to include all of these steps. Correctly designed and implemented, a dashboard has the potential to bring immediate and considerable return on investment (ROI) to your organization.

1. **Plan:** Dashboard development begins in the planning phase. Identify the project team members, their roles and overall project objectives. When working within a tight timeline, populating the dashboard is the most critical area of concern. Take care not to underestimate the complexity of the databases in which the data resides. Accessing the data from a myriad of tables requires technical resources with detailed knowledge of the underlying table structure and considerable SQL skill. Define the project budget and take into consideration the work required to create the custom queries for the desired metrics. Set realistic goals for your dashboard project by striking a balance between the primary user’s needs and what you can afford to deliver.

   **Requirements gathering and prototype**—Interview the key stakeholders to determine their needs and expectations. To keep the dashboard project within scope, map these needs and expectations to the pre-established KPIs. To increase the likelihood that the final dashboard will meet users’ expectations, take advantage of available tools and technologies that lend themselves well to prototyping.

   **2. Design:** Once the team approves the dashboard’s content and appearance, the next step is to incorporate major design aspects:
   - Refine the user interface and control flow;
   - Confirm the data sources for each data element;

   “It may seem relatively easy to build a dashboard. However, such oversimplification can lead to a failed project.”

   3. **Build and validate:** The “real” development begins at this stage of the project. Several tasks occur here, typically in parallel and closely coordinated with each other.

      **Front-end implementation**—Create the dashboard user interface. Evaluate what graph and chart types best represent the data to be displayed and make decisions regarding grouping data to provide the greatest visibility for cross-analysis.

      **Query implementation**—Create the queries to retrieve the necessary information from the appropriate databases. This step can be particularly complex and time consuming, especially if there are multiple data sources for the various dashboard elements including data from customized enterprise applications for ERP, CRM or SCM.

      **Configure scheduling, refresh and security**—To ensure the dashboard content is up to date, queries need to be configured to run regularly. At the same time, it is important to establish and implement security rules to display the appropriate information for users with different levels of access.

      **Dashboard validation**—As with any software project, when the effort reaches “code complete,” both the technical team and the primary users must test the dashboard to ensure it meets the requirements outlined in the project plan.

   4. **Deploy and maintain:** Once the dashboard has been built and tested, it is then deployed into production and security requirements are implemented. With the
Florida Rock Institutes Customer Self-Service Dashboards

Founded in 1945, Florida Rock Industries, Inc., is one of the nation's leading producers of construction aggregates, ready-mixed concrete, concrete block, Portland cement and pre-stressed concrete. The company is also at the forefront of technology in the construction materials industry.

When Florida Rock receives a customer order, it assigns a “ticket” to each request for materials. The ticket acts as the proof of delivery. It is signed by the customer when the shipment is received and scanned into Florida Rock’s electronic document retrieval system. Once the signed ticket enters the system, Florida Rock adds the associated customer order to an invoice. There can be one ticket or a hundred tickets on an invoice, so tying the ticket back to the invoice quickly and efficiently is critical for providing the kind of quality customer service for which Florida Rock is known. The company can field hundreds of calls per week from its more than 8,000 customers looking for information regarding their invoices. The process involved a customer going through a telephone answering service to reach the correct person, who would then query an Oracle database or Florida Rock’s electronic document retrieval system to get to the needed information. This image or copy would then be emailed, faxed or mailed to the customer. Florida Rock also saw the need to make it more efficient both for its customers and for its credit and accounts receivables (AR) departments, who spent several hours a day just supplying lost copies of documents.

**Addressing the Challenge**

The solution involved a creative way to use the Noetix Dashboard as a customer self-service tool. Instead of having to call for information, customers are able to sign on to a secure system that takes them directly to a custom-tailored Noetix Dashboard. It shows them their AR buckets; allows them to view their most recent 50 open invoices and all the associated tickets, and lets them see receipts or checks that have been posted to their accounts. Finally, it gives them a list of local Florida Rock locations if they need to do more business. Essentially, Florida Rock’s customers are getting a great deal of information about their business in a matter of seconds.

In addition, customers can link directly from their invoices to Florida Rock’s electronic document retrieval system and pull out a PDF with the exact image of the document, or drill down into the tickets that made up that shipment or particular invoice, as well as make a copy of their most current statement.

The Noetix Dashboard provides Florida Rock customers with a self-service tool to facilitate a more efficient business partnership. It is much easier for them to monitor their accounts and ensure their invoices are paid on time, cutting back on overdue payments. “Our customers require that all of the signed tickets be filed,” said Dave DeVore, manager, application development, Florida Rock. “If one ticket is misplaced and there is no proof of delivery, payment is withheld on an invoice that could be worth several thousand dollars. The easy-to-use Noetix Dashboard will help shorten the payment cycle by giving customers immediate access to all of the tickets and invoices associated with each account. We want our customers to have control over their businesses and get the information they’re after in a timely manner. Logging into these dashboards is certainly a lot quicker and easier than calling, going through the various levels, and then waiting for a fax or image to come through.”

In addition to the customer self-service tool, Florida Rock uses Noetix Dashboard internally. The company’s credit department staff uses a dashboard to get quick customer snapshots and access to associated statements, invoices and tickets without having to go into Oracle ERP. Florida Rock also intends to roll out dashboards to its executives in HR payroll and receivables to monitor business operations from a single, intuitive view.

**Final Note: Build vs. Buy**

Building and deploying an executive dashboard takes time, regardless of the vendor or technology that is chosen. Creating the graphical front end is relatively quick and easy, but that’s merely the shell of the dashboard. What you actually see on your desktop pales in comparison to the hidden effort—80% of the complexity lies beneath the surface. All of the tasks listed above require planning, organization, coordination, scheduling and solid project management. When comparing proposals or considering a “build vs. buy” decision for deploying a dashboard, it is imperative to ensure that the entire scope of the project is considered.

Noetix provides business intelligence tools that enable more than 1,300 customers worldwide to quickly and cost-effectively access the enterprise application data they need to make important business decisions. Unlike most BI tools that require weeks of extensive manual mapping to be set up and maintained, Noetix uses patented technology to automatically discover and produce metadata based on customers’ specific implementations of Oracle Applications or Siebel CRM. For more information please visit www.noetix.com.
A Sustainable Advantage

Building the Search Center of Excellence

By Hadley Reynolds, VP Center for Search Innovation and Silvija Seres, VP Strategic Market Development, FAST

Search is strategic; however, the strategic potential of search is not captured by the act of acquiring a powerful search platform alone. Pioneering firms are now developing a new kind of management approach to help deliver maximum value across multiple search-driven applications: the “search center of excellence.” It is a structured approach, utilizing a focused cross-functional team, and it is emerging as a practical tool to drive search innovation and deliver high quality online experiences.

This is the age of search; search is becoming the de facto infrastructure for finding and delivering information. It is ubiquitous in new online business applications, driving revenue and capturing operational efficiencies inside the organization. Any organization whose operations touch the Internet, or important digital information in general, is finding that delivering better search is good for business.

Yet despite the scale and importance of this trend, many companies can’t seem to get out of their own way as they begin using search. For example, many firms have fallen into what we refer to as the “one-size-fits-all” technology purchase syndrome. In this mode, the enterprise search problem is seen (at least by the sponsors) as solved as soon as new “enterprise” software is installed on a production server. In such cases, however, the value of the solution often fails to impress users inside the company or customers and partners outside, because it simply does not seem to “get” their particular business situation. This is because the core of all successful search experiences is built on understanding the enterprise business context and the knowledge drivers that power each specific set of business interactions.

One indicator of the challenges posed by this current state of the practice is that virtually all researchers into search quality continue to report user frustration in both external and internal applications of search. Forrester Research, for example, has consistently reported breakdowns in site-search quality. Recent research shows that 58% of 211 websites reviewed through mid-2006 failed to meet basic criteria for site search engine and search interface quality. Failure rates for clarity and presentation of navigation options were in the same range or higher. At the same time, the firm’s demographic research finds that findability and navigation are even more important to online site visitors than the quality of information on the site or the range of functions available.

Looking at information breakdown inside the organization, IDC Research has found consistently that the cost of wasted time on the part of professionals searching but not finding information is a major continuing cost to organizations. The most recent 2006 “Hidden Costs of Information Work” report suggests that this cost amounts to $5.3 million annually for an enterprise with 1,000 information workers. Delphi Group reported that more than 50% of professionals surveyed report being either dissatisfied or very dissatisfied with the search experience in their firms, while only 15% reported that their firms had an enterprise search strategy in place.

These research examples show that the issues with achieving high-quality search go much deeper than a selection of technology. Some of the problem is clearly related to the legacy of basic keyword search deployments whose fatal lack of accuracy and relevance continues to disappoint users. Most modern platform offerings combine families of advanced linguistic and statistical functions that are more than adequate to deliver highly accurate and contextually significant results in a rich analytic framework with suggestive and intuitive navigation options. We maintain that many of the issues of search quality arise not from technology limitations, but from the unnecessarily limited implementation practices which most firms have resorted to in deploying search.

Fortunately there is a constructive solution to the challenges described above; we are seeing impressive results at a number of firms who are making search quality a priority. Businesses as diverse as Merrill Lynch, Pfizer, McGraw-Hill, Autotrader.com and YouTube are dramatically raising the level of the search experience they offer their online audiences. They are replacing “one-size-fits-all” thinking with a management process that secures business acumen and measured investment strategies at the center of the search deployment. The new focus is on developing core organizational resources and tailored governance capabilities that will deliver business value across multiple search-powered applications.

Search Quality Drivers

Before we describe the approach in more detail, let’s take a closer look at the kinds of competitive and business drivers that lead these pioneer firms to deliver best-in-class search.

The Internet has made everyone more demanding when it comes to search performance and intelligence. Customers and employees have all become acclimated to the apparent effortlessness of Web search on MSN or Yahoo! or Google.

Self-service is no longer just for shops or gasoline “service stations.” Today it is also the accepted access model for information. Customers and employees now require, as well as expect, self-service tools able to mine all the information sources they should have access to and to deliver relevant results in a familiar and comfortable environment.

Organizations that can deliver the right information at the right time with the right search behavior reap dividends from increased online sales and from empowered employees. In order to gain these returns, information access needs to go beyond the “one box/one button” paradigm and adjust user experiences to match their

Many firms have fallen into the ‘one-size-fits-all’ technology purchase syndrome.”

58 KMWorld November/December 2006
Typical Roles in a Search Project

1. Executive sponsor—budgets and overall success of project.
2. Business owner—representation of business/user community and clear definition and communication of their requirements to the project.
3. Program/project manager—all project planning, resources, communications and deliverables.
4. Information architect—content planning and management (e.g. meta tags and taxonomies).
5. User interface engineers—design, development and integration of search front end with existing applications.
6. Hardware engineer—all hardware and O/S installs, in addition to DNS, DB or other software.
7. Network engineer—network configurations as required for the implementation.
8. Operations—daily operations of search solution, including all first-line support.

The COE practice brings together people with deep business-domain expertise, broad search-applications experience, cutting-edge software infrastructure knowledge, complex project-management skills and demonstrated facility in knowledge transfer. This group has the ability to act as a central point of contact to facilitate collaboration between lines of business, functional specialties and customer, service provider or partner resources. It may provide the resources to staff each of these components of search projects: application architecture and design, project methodology, best practices and standards, user interface design, education programs, support services and analytics for continuous improvement.

By putting a dedicated team in place, companies adopting the COE process gain the ability to:

◆ Identify core patterns of search success;
◆ Share best practices and facilitate innovation in next practices; and
◆ Leverage search technology, knowledge engineering and search infrastructure skills across the enterprise.

The Search Center of Excellence Practice

Using the COE to create a repeatable process, common business rules, standard best practices and custom methods and components tailored to the business will drive down the cost and improve the success rate of implementing search projects.

By providing a project office for search, the COE practice can integrate training programs, business consulting and project portfolio prioritization, best practices examples, advanced solutions “tiger teams,” implementation services, application monitoring and continuous improvement services. These capabilities deployed within the context of an organization’s business priorities offer a fast-track approach to high search quality. In our view, it’s time to adopt this approach to driving business innovation with quality search.

Hadley Reynolds works in Boston as a vice president and the director of the Center for Search Innovation at FAST. He has been an industry thought leader for over a decade, researching, speaking and writing on new trends in search technology and business practice.

Sibija Seres works in Oslo as a vice president of strategic market development with FAST. She holds an MBA from INSEAD in France and has extensive scientific background in algorithm design and optimization, with a Ph.D. and Prize Fellowship from Oxford University.

Enter the Search Center of Excellence

In adopting a “search center of excellence” (COE) approach, organizations are moving search from the level of technology detail to the level of business innovation and strategy. With a strategic approach and executive-level support, they are adopting a centralized management capability to turbocharge a diversity of search projects across the enterprise.

The development process for the COE takes place on a number of levels. In order for the COE to succeed in its management and strategy dimensions, it must first and foremost gather sponsorship and authority at the senior-executive level. The COE governance model establishes the interaction protocols between the COE and the various business units and technology support groups of the enterprise. In order to align search projects with business objectives and to leverage the benefits of knowledge sharing, experiential learning and technical search expertise, the COE unifies executive-level interactions among all the business units who are or will be making use of search technologies.

Within the operations of the COE itself, the key success factors are: (1.) ensuring that appropriate roles are identified to support the anticipated activities of COE projects (see sidebar, “Typical Roles in a Search Project”) and (2.) ensuring that the competency models and interaction patterns for those roles are thoughtfully specified.

Search technology makes unique demands across the entire spectrum of traditional IT roles, from systems analyst to architect to developer to database administrator to user-interface designer. It also introduces non-traditional knowledge engineering and customer-experience management elements to projects and programs.

The COE leverages resources from across the firm and potentially across the customer and partner universes. In order to deliver high business value and quality user experiences, the COE incorporates input and participation from line-of-business managers, business-process designers and business analysts, human-factors experts, business-intelligence analysts, merchandisers, marketers and other stakeholders of search applications.

Roles, the context of their questions, their vocabulary and their purchase or work patterns. Users do not want to know about the multitude of different formats being consumed, analyzed, contextualized and personalized for consumption by the search platform—they just want the system “to work” across this universe of information, with the most suggestive and relevant results. The best of these results are, in fact, delivered by composite and intelligent business applications, built on the search platform and engineered to unify views of an arbitrarily complex information “space.”

With this new class of intelligent business applications, “search” is moving far “beyond the box” and taking up a role that is as central to today’s Internet-connected businesses as relational databases and ERP systems were to the pre-Web era of IT. In the most advanced enterprise and commerce implementations today, “search” acts as the crucial business information filter—bringing to life the “long tail” value resident in both enterprise and Internet information.

Taking advantage of these new capabilities in search platforms requires an evolution in internal processes and practices for aligning business goals with technology management. It requires a new collection of skill sets for developing composite applications with disparate data. It requires a change in thinking from data models to consumption paradigms.
Implementing KM

Practitioners Share Best Practices

By InQuira

M ost customer service organizations today will admit that effective knowledge transfer is the most crucial element to resolving customer problems. When done correctly, knowledge transfer accelerates problem resolution processes, fuels customer satisfaction and leads to greater organizational efficiency. Organizations invest in processes and technologies that enable them to create, manage and publish knowledge, and that allow them to find, retrieve and share the enterprise’s knowledge across all support channels. Without formal knowledge management processes, companies would be unable to share knowledge with their customers, partners and employees.

When organizations consider investing in an enterprise-level knowledge management (KM) initiative, they generally conduct a detailed analysis of how to effectively tackle a project of such magnitude. Companies want to know how to best prepare for such projects, what pitfalls to avoid, and how to measure results that are generated from their KM implementation. They scrutinize internal short-term and long-term needs, organizational objectives and resources, customer needs, system requirements and technology implications. Companies conduct knowledge assessments, system assessments and develop a roadmap by which to guide the initiative. They consider the role of and impact on stakeholders, end users, partners, internal teams, customers and outside vendors. They try to anticipate and quantify the impact of the knowledge management project in their business processes, operations and ROI.

No amount of analysis and planning, however, will uncover every potential challenge or roadblock. Here, practitioners who have been through successful large-scale knowledge management initiatives share their insights and lessons learned.

Best Practices

1. Over-communicate. “You can never underestimate the change management requirements that go along with projects like these,” according to Jodi McBride, director of knowledge and training services at Pitney Bowes. “You need a shared vision, supported and driven throughout the company from the highest levels of the organization.” Communicate regularly throughout all the phases of the project with business stakeholders and end-users. Follow a communication plan that keeps all parties included and responding. “You are going to be taking people outside of their comfort zone and they need to understand what you are doing. You can never over-communicate with a project like this,” said McBride.

2. Recruit the right people for the project. Identify and recruit the right people with the right skill sets from the start, making sure their skill sets align with the assigned tasks and responsibilities. Maintain a consistent vision throughout the implementation project. A knowledge manager at a leading mobile communications company suggests that practitioners ask themselves “whether the people who defined the specifications have been involved all the way through the implementation. Sometimes as people change in a project, it is easy to lose sight of what you originally thought versus what the new folks thought as they were introduced into the project.”

3. Solicit end-user input in the solution design. Involve end users in content identification, design and testing. Consider involving different representatives from different departments. “At Pitney Bowes, we created a role called ‘content ambassadors,’ which included representatives from different departments within the contact centers. A content ambassador’s job was to help identify sources of content; validate, enter and test content; help design the user experience; and provide feedback on which information should be made public for employees and customers. Without a doubt, their contributions led to greater user adoption,” said Pitney Bowes’ McBride.

4. Encourage user adoption with incentives. Employees may be reluctant to move out of their comfort zones or embrace new processes. Help employees understand the desired results and how they will be measured. Then, develop ways to recognize and reward individuals who adapt to the changes in the system to encourage its use.

“Rewarding and recognizing employee efforts to share knowledge is a powerful way to encourage this activity. Consider periodic bonuses to individuals with exemplary efforts, peer recognition combined with cash, and project work available to individuals who actively share knowledge as three incentive strategies,” suggests Ladd Bodem, principal and co-founder of market research firm ServiceXRG.

5. Identify content gaps and duplicates, and scrub your content early. Eliminate redundancy and pieces of content that overlap. Although content scrubbing is an ongoing process, more scrubbing in the early phases of the KM project can help make content more usable and help improve user adoption since users are able to quickly locate the right content and do not have to go through duplicate content.

“Our content has always been organized by department,” commented McBride. “For the first time, a centralized knowledgebase provided visibility to content overlaps between teams. Additional time to test and scrub responses that overlapped would have been beneficial in reducing the amount of redundant or conflicting content.”

6. Define ROI measurement and reporting requirements early in the process. McBride emphasizes, “Make sure you understand exactly what you are going to look for to get to your ROI and define those reporting requirements as early as you possibly can.” Once live with the new system and process, expect business stakeholders to ask for reports comparing performance to expected ROI from the project.

According to a recently published report by ServiceXRG titled “Knowledge Management—Strategies. Benchmarks and Best Practice,” “there are two distinct types of knowledge management measurements; one type is focused on measuring the efficiency of the KM processes such as content coverage and quality content, while the other looks at the impact from knowledge management. Some impact metrics to consider include: measuring deflection, staffing, change in first contact resolution rates, resolution times, agent productivity and overall success rate.”

7. Don’t underestimate the impact of tangential benefits. For example, Jodi McBride explains how at Pitney Bowes, “we didn’t anticipate how, having information available from a single source, was going to change how we trained. Our focus this year has been on redesigning the classroom experience to incorporate the use of KIP (Pitney Bowes’ knowledgebase) and create more interactive training materials that better engage the learner across multiple dimensions. This is a dramatic shift away from typical lecture-based training programs. This shift in the classroom is not something we anticipated when we went live, but has been a positive experience nonetheless.”

8. Know the needs of your end users. Understand how language components impact search accuracy. Identify user search behaviors and consider search rules to improve the user’s experience. “Understand how people use the resources they have available today. Understand the slang for key concepts, and how people from different departments may approach the same information but from different perspectives. These
About InQuira

The InQuira 8 suite of business applications improves companies’ Web self-service and assisted-service resolution processes. With applications for knowledge management, Web self-service, agent-assisted service, email deflection and customer experience management and analysis, InQuira 8 delivers a complete solution for resolving customer support issues online and in the contact center—applications that can infer the user’s needs and intentions and provide the right information, tools, recommendations and assistance in an efficient, orchestrated interaction that will resolve more issues and deliver greater satisfaction and higher ROI.

**Knowledge management.** Many customer support issues require research to resolve, either by the agent in the contact center, or by the customer in the self-service channel. A core function of support organizations is to capture enterprise knowledge to create content designed to resolve customer issues. InQuira 8 provides full knowledge management capabilities, including:

- Capturing content and content requests from within the resolution process;
- Authoring process that removes the unnecessary tagging and increases content reuse;
- Full version control, with the ability to revert back to a previous version, compare versions or view versions side-by-side;
- Robust publishing workflow to ensure content is effectively managed through user-defined stages of development and publication;
- Task management to ensure that each task is being completed by someone with the right role and skill;
- Configurable email notifications for all tasks in the system directing attention to complete work;
- Multi-lingual content and translation workflow to update and distribute content in multiple languages;
- Embedded natural language search (InQuira Intelligent Search) for more effective and useful retrieval of information from within the knowledgebase and forum content;
- Discussion forums for community-based support;
- Embedded reputation models to support performance evaluation for knowledge engineers;
- Tokenization to secure content and sections of content for specific user audiences; and
- Configurable subscriptions to categories of content and specific content items, including newsletters and content from specific authors.

**Web self service**. Customers and companies increasingly demand more sophisticated and effective Web self-service capabilities to resolve customer support issues. InQuira 8 delivers that. Key capabilities include:

- Unique ability to determine user intent from search and navigation behavior, and use that insight to manage the customer experience intent-by-intent, empowering companies to deliver a personalized resolution experience based on an understanding of what the customer is trying to accomplish;
- Packaged horizontal and verticalized Web applications to accelerate implementation;
- Industry-specific “intent libraries” to map the resolution experience to pre-defined customer needs;
- Intelligent search and retrieval capabilities based on patented natural language processing technology;
- Industry dictionaries to improve search effectiveness; and
- Diagnostic process wizards to troubleshoot and resolve customer problems.

**Agent-assisted service.** Contact center managers are under constant pressure to deliver effective, loyalty-inducing customer service at reasonable cost. InQuira 8 for agent-assisted service includes the following productivity-enhancing capabilities:

- Embedded intelligent search, retrieval and navigation from within the agent cockpit;
- Integration into leading CRM packages, including Siebel and Clarify;
- Ability to associate retrieved content to case resolution;
- Web self-service session history captured on service requests escalated to agents via email;
- Integrated links to trigger knowledge creation workflows from specific cases; and
- Collaboration.

**Email deflection.** When customers are unable to resolve their own problems from a company’s website, they will often submit their questions to the company through email. InQuira 8 for email deflection:

- Intercepts Web-submitted emails and uses the subject line to search and retrieve appropriate information to resolve the customer issue;
- Deflects inbound service requests when a customer clicks through the offered solutions and abandons the submission process;
- Passes service requests directly into the ERMS when an appropriate solution is not found; and
- Captures Web self-service session history and passes it with the service request to the agent in the contact center, allowing the agent to pick up where the customer left off.

**Experience management and analysis.** Companies recognize that customer service is an iterative business process based on continuous analysis and refinement. InQuira 8 offers several capabilities to manage the customer experience and measure the effectiveness of the resolution process and the creation and use of knowledge assets. Capabilities include:

- Intent libraries for retail banking, telecom and automotive industries that automate the mapping of searches and navigation behavior to industry intents. InQuira enables targeted and managed responses at the intent level, resulting in a 10:1 reduction in the number of managed responses users would have to define;
- The ability to guide an interaction and present additional content and offers that result in higher resolution rates; and
- Out-of-the-box data warehouse, star schema and reports to analyze user interactions. InQuira Analytics includes full ability to drill into reports to look at relationships of data for in-depth analysis of results. The analysis that can be performed using InQuira analytics includes conversion analysis, case escalation and email ROI, navigation usage, content gap analysis, content usage, content authoring, customer feedback, surveys and process wizard usage.

9. Keep the end goal clearly in mind. Periodically evaluate how well you are adhering to your original specifications. It is important to make sure that decisions on whether to expand the effort or to maintain the effort are based on your business needs and the timing you are trying to achieve.

10. Consider the impact on existing business processes. Identify and understand all the existing processes that will no longer work with the new system so that these can be addressed. A knowledge manager of a leading mobile communications company notes that, “you need to understand the processes that will change and what is going to break or be different as you go into the implementation phase of your investigation.”
Knowledge Management through Portals

Mistakes to Avoid and Principles for Success

By Brandon Lackey, Global Solutions Director, BEA and Michael Behounek, Managing Partner, Emerja LLC

Many organizations today still struggle to get value from their knowledge management (KM) efforts. Even companies that have been benchmarked as the “best practice” can easily fall from the top. So, how can you get sustained value from KM initiatives? There are several key factors for valuable KM deployments. But first, let’s learn from what has not worked. Here are the eight most common KM mistakes.

KM Mistakes to Avoid

1. Build something and hope that they’ll come. This could be said for almost any technology project, but it’s particularly common in KM. Why? Management is looking for the quick-fix, a silver-bullet technology. It’s faster and easier not to get the users involved. A global mid-sized company embarked on a project to help the organization efficiently match expertise with their customer support needs. Their IT group built an expert locator system that was implemented using a traditional change management approach. In a project status report to management, three months later, they found out only 72 out of the 2,500 target users had completed their profiles. Six months of desk-pounding by management got the total up to over 200 users. Subject matter experts claimed it provided no value, and it took too much time to update. It was also not tied to any other data sources that could be leveraged to make the users’ jobs easier and to help them keep the information updated. It was not until a portal was implemented that the system began to reach critical mass.

2. Implement technology tools without a business owner or a specific business problem. A Fortune 100 company felt it had a reasonable understanding of the business problem. A project team was commissioned to produce a new enterprise framework and a pilot for a community of practice. In the design phase, the team realized that the scope and the expected target users were off mark. So they adjusted the design. One month later, the original technology cost assumption was found to be off, and the original business justification was now ambiguous. The team began to build a new business case with a viable ROI but a different business sponsor was now required. The critical business owner declined to participate and the project was eventually cancelled. If a business case had been done before starting the project, the correct audience, business owner(s) and technology could have been anticipated.

“If a business case had been done before starting, the correct audience, business owners and technology could have been anticipated.”

3. Deploy KM Enterprisewide. Knowledge in most situations is contextual. Implementing an approach across the entire enterprise poses significant risk of failure. According to research by Robert Francis Group, up to 85% of all KM initiatives fail to achieve their business objectives. For example, a large energy company’s IT group was asked to identify a technology solution for collaboration, in both teams and large groups, as part of the KM effort. Several technologies were evaluated and one was selected for the 50,000 enterprise users. Although the technology was best-of-breed, the standalone application was deployed to the entire enterprise. Adoption was slow. The implementation team did not answer the fundamental user question: “What’s in this for me?” The result was a costly re-launch. This time, however, it was done on a project basis around core business groups with critical business challenges. In addition, the stand-alone application with other project-specific applications was surfaced on a portal platform. User adoption rate more than doubled.

4. Assume knowledge is only in documents and data. If “content is king,” then the king is dead. Most companies are drowning in content and most of it is irrelevant. The most valuable knowledge in an organization is “implicit knowledge.” As part of a KM initiative, a VP of an engineering firm launched a project to update and maintain engineering standards. Believing that this explicit knowledge was the key to knowledge longevity and transferability, the firm bought and installed a content management system. The system suffered from low adoption. Further exploration revealed that the less-experienced engineers had trouble applying the information to their specific problems, and the more experienced engineers only rarely needed it. A better solution would be to provide a continual method of capturing real-time knowledge from threaded discussions or instant messages, and indexing the salient points for quick retrieval and evaluation. Implicit knowledge would then be a result of the collaboration and made explicit and usable in context.

5. Not integrating unstructured content with the structured data. Since most companies have terabytes of unstructured content, they rely on search engines to index the documents and users to find the information by keywords. Unfortunately, the user is often unaware of additional structured data that relates to this unstructured content. Most documents do not contain the unique identifier that unlocks the door to the databases, applications and systems that contain a wealth of other relevant content. Most deployments do not take into account the relationships between the data and applications. Portals can help bridge this gap more effectively by linking disparate data and content sources through an interface to seamlessly present the aggregated results.

KMWorld November/December 2006
“Executive support is critical since it involves the assignment of people and a change in process.”

KM Principles for Success

Avoiding these eight common mistakes can help you reduce your risk of failure in KM. But what should you focus on? Results for the organization can be achieved following these simple practices:

Solve real business problems. This may sound like a no-brainer, but in the throes of a technology roll-out, managers may lose track of the underlying business problems to be solved. Clearly stating the problems has further benefits in scoping the requirements, managing deliverables and selecting quantifiable metrics. In addition, objectives and benefits that are clear to users have a positive effect on the adoption process.

Secure executive commitment to KM. Executive support is critical since it involves the assignment of people and a change in process. Gaining long-term commitment from executive sponsors comes from solving their business challenges, delivering business value and, ultimately, gaining their trust. Keep them informed and have them play a visible leadership role in the launch.

Own the process within a business unit or organization. Unlike many support functions, KM needs to be embedded in the business unit or support function. Having it “outside” the groups delivering the value will only hamper widespread use. A small KM core team within the business unit is required to help facilitate the solution and help ensure similar processes and scalability between projects.

Emphasize people and process. Unfortunately, in many cases, projects are driven by the technology. One secret to getting people to participate and collaborate is to ensure they get value from doing so.

This can take many forms—streamline business processes, reduce their workload or enable them to spend time on job-related initiatives with people inside and outside the organization.

Integrate the process into the employees’ workflow. Weave KM processes into the normal workflows and, where possible, simplify and optimize for the particular user. If the solution is extra work or a bolt-on, user adoption will remain low.

Capture business metrics to determine success, drive improvement, and communicate your achievement. Too often, the only captured metrics are for portal or application usage. Executives and users want to see value that relates to the business proposition, which can involve tools like business activity monitoring and analytics. For example, a portal community of practice to support technician repairs would track and see improvement in failed repairs, average repair times and repairs per technician with high portal site usage. Having limited or no business metrics weakens justification during the budget cycle or when a change in management requires a review of the project value.

We learn as much from our mistakes as from our successes. The challenge is to take what others have learned, to avoid the same mistakes and to produce a consistent, repeatable KM program that delivers significant business value.

Brandon Lackey is global solutions director at BEA. He assists organizations in aligning business objectives with their enterprise portal strategies. Before joining BEA, Lackey led the development of Halliburton’s worldwide portal that grew to serve more that 6,000 clients and 200,000 users. During Lackey’s tenure, that portal received awards from Network World, the DM Review and InfoWorld. Contact: brandon.lackey@bea.com.

Michael Behounek is Emerja LLC’s managing partner. Before forming Emerja, Behounek had 25 years in the energy industry. In 2001, he led the development of Halliburton’s highly successful KM effort that resulted in a 563% return over a three-year period. He continues to help companies deliver breakthrough performance by using collaboration to solve complex business problems. Contact: michael.behounek@emerga.com.
Knowledge is power. Though said in the 16th century by author and philosopher, Sir Francis Bacon, the meaning of the phrase is even more applicable today as companies around the world are dealing with an overwhelming amount of data stored on computers, servers, flash drives, PDAs and other storage devices scattered throughout the enterprise. The information, email messages, documents, spreadsheets, presentations, graphs, photographs and other content trapped in information silos across the enterprise, contain a significant amount of knowledge that can benefit the organization. Unfortunately for many companies, much of this information remains untapped, unsearchable and trapped in disjointed systems.

Enterprises have deployed a complex web of information management systems, applications, operating systems and databases, to manage and deploy data and content. However, the resulting information-access systems are unable to exchange information in an efficient, dynamic fashion because they lack a common understanding of the business semantics that describe the information contained in the various content silos. Marketing may have one way to describe a new product, while IT, shipping, public relations and research and development may each have others. When something in the product lifecycle changes—a name, a feature, a color or price, for example—then searching and finding all the information across the enterprise related to that product can be a chore, if not impossible. Business semantics management provides organizations with a dynamic registry to model, govern, publish and collaborate around a consensus of terms and term relationships that define the products, services and organizational knowledge of the enterprise.

Business semantics management is emerging as a core component of service-oriented architecture (SOA), which distributes applications that perform services on demand. For many experts, SOA is the future of software, where applications, used internally or externally, are delivered as services. Companies such as SchemaLogic of Kirkland, WA, have developed a business semantics management solution that models complex business semantics relationships and provides those models to subscribing systems as a service within the context of an SOA. Referred to as “semantics-as-a-service,” the solution provides a single source for creating, managing and distributing corporate semantics.

Semantics-as-a-service can be delivered as a real-time service to enable organizations to instantly update their semantics models as changes happen, or manage their semantics with regular and consistent updates. Semantics-as-a-service can be deployed behind the firewall or as a Web-based application. As enterprises move towards SOA and Web services environments, semantics-as-a-service has emerged as a business-critical solution that can help to enhance knowledge management. The evolution of semantics management from a manual time-consuming, cumbersome process to real-time through semantics-as-a-service is a revolutionary concept that is changing the way companies are managing their information assets and their businesses.

Increasingly, companies are finding that developing a truly effective enterprise lexicon requires the participation of all stakeholders in a company, from the purchasing manager to the CTO. Semantics-as-a-service allows organizations to collaborate and leverage the knowledge and expertise of a variety of sources, including subject-matter experts, business users and domain experts to develop a semantics model that describes the corporate lexicon, provides up-to-the minute know-how and makes knowledge accessible across the organization. SchemaLogic’s centrally managed solution and Web-based collaboration service helps to develop corporate semantics models that can be used enterprisewide to describe their business, products, services and overall corporate knowledge.
“Information-access systems are unable to exchange information because they lack a common understanding of the business semantics.”

Kellysearch: Managing Search

While IBM focused on using business semantics management to manage its collective knowledge about its human resources, Kellysearch, the largest business-to-business (B2B) search portal in Europe, was faced with a growing taxonomy that its infrastructure could not adequately manage.

The company’s growing taxonomy included more than 200,000 phrases and 400,000 sub-phrases that were attached to more than 2 million businesses. Unfortunately, the company’s current infrastructure and manual taxonomy processes were not capable of scaling to manage the growing amounts of data, which resulted in an increase in time-to-market for developing and publishing advertiser information to the Web.

Kellysearch, a division of Reed Elsevier, needed a knowledge management solution that could drive queue time down and improve the speed and quality of search results in order to increase market share and the revenue stream of its online model. Kellysearch turned to the semantics-as-a-service model to link search with content to drive revenues and move the classical advertising print business to the paid search market on the Internet.

Kellysearch now has the capability to establish and build corporate semantics by initially harvesting data from Kellysearch’s multiple company and product listings. The information is then rationalized so that it is established, evolved and distributed across Kellysearch. The variances in naming terminology among companies are quickly resolved to a common set of terminology, making it easier to match global B2B queries for specific goods and services. By implementing a business semantics solution, embedded knowledge is monetized via an improved search and ad-revenue model. Kellysearch now has the ability to quickly expand internationally to accelerate revenue; business processes have been streamlined and content contributors empowered; and enhanced search results have increased the quantity and quality of leads to paid advertisers.

In both the IBM and Kellysearch examples, information that has typically been difficult to manage or out of reach, is now available to everyone, driven by semantics-as-a-service. Delivered through an SOA framework, semantics-as-a-service provides organizations with the flexibility and scalability to meet the demands of their growing businesses, manage the contributions from all areas of the enterprise and increase the speed of change.

Knowledge can be a variety of things to different companies. For some, it’s the magic sauce or company trade secrets; for others it is the intellectual property found in approved patents. Today’s enterprises need a common semantics model to manage information across the enterprise. Semantics-as-a-service is emerging as a business-critical application, and for companies like IBM and Kellysearch, has become a mission-critical application for their services-oriented architectures.

More than 400 years from its origin, the phrase “knowledge is power” couldn’t be truer. The ability to unlock the corporate assets of human capital and products for competitive advantage; to make business decisions based on improved data quality, collaboration and agility; and to better manage information and knowledge for the benefit of the entire enterprise is true power.
Answering Customers’ Questions the Intelligent Way

The Importance of Multiple Search Methodologies

By Anand Chopra, Director, Product Marketing, KANA

Enterprises face a difficult challenge when it comes to simultaneously improving the quality of customer service and reducing service costs. More products, growing product complexity and rapid change substantially increase the amounts of information required to answer customer questions and troubleshoot problems. Paradoxically, this growth of information availability increases the difficulty of finding relevant solutions.

For enterprises to improve self-service adoption rates, increase call center efficiency and improve response accuracy, they need solutions that help agents, customers, partners and suppliers find answers more efficiently. The traditional methods of search and retrieval use keyword, simple text and Natural Language Query (NLQ). In many cases, search-based knowledge management solutions emphasize their ability to sift through multiple enterprise systems to present results. This typically generates long hit lists with many irrelevant entries. Another failing of this approach is that results are presented indiscriminately. The user does not know if the information is accurate or current, increasing the possibility of an incorrect or out-of-date answer. Unfortunately, these methods are best suited to expert users who are familiar with the content and terminology and know which words will most quickly yield a correct answer. Novice users without domain expertise cannot easily apply the terminology precision these techniques require and, frequently, people need guidance to find the answers to a question.

Call center agents, customers and partners can all benefit from knowledge management solutions that organize and structure access to information, with intelligent guidance that matches each user’s level of sophistication and skill. In addition to keyword, text and NLQ search, a knowledgebase should deliver a set of sophisticated search and retrieval methodologies that guide users through the issue resolution process so first-time self-service users and highly experienced agents can quickly find the right information including:

- **Case-based reasoning** combines NLQ with clarifying questions. The user enters a text string that yields a solution set and a series of targeted questions to further narrow the problem definition. Based on the user’s answers, case-based reasoning narrows and reorders the solutions in order of relevance.

- **Decision trees** guide users through structured diagnostic scripts. Each time the user answers a question, the decision tree dynamically presents new questions and narrows down the number of possible solutions until the most appropriate solution is identified.

- **Expert modeling** ranks potential solutions in order of relevance to the problem as determined by subject matter experts. Expert models define precise relationships between problems, causes and solutions.

Using this wide range of search and retrieval capabilities, enterprises can offer multiple levels of guidance and allow users to select those techniques that best match their skills and preferences. However, the search methodologies should be blended seamlessly so that users do not have to make a conscious decision about which method to use. For example, users can begin a session with a text string search, and then be presented with a series of questions that combine aspects of case-based reasoning, decision trees and expert modeling to quickly refine the general description into a specific problem description. KANA customer John Harrigan of Siemens states, “What attracted us was that these are the kind of tools that you can get the average user up to speed on without requiring much background knowledge of the actual system.”

Reporting and tracking capabilities should complement these methodologies to dynamically score and rank potential solutions by popularity based on users’ experiences. Users provide feedback with each search on the helpfulness and accuracy of the solution, which is incorporated into future search results so that solutions are scored higher or lower on subsequent similar queries. Paul Kinsella, VP worldwide customer response at Creative agrees, stating, “with the knowledgebase feedback mechanism, we can update and revise content to reflect customer choices and preferences.”

As the management of and access to knowledge grows in importance, so does the ability to deploy a comprehensive, yet maintainable knowledge management solution that increases customer acceptance of self-service, enables call center agents to answer inquiries more quickly, accurately and consistently and enhances the value and use of information stored in enterprise systems. By providing multiple search methodologies, enterprises can empower users of all levels to diagnose and resolve problems with ease.

Anand Chopra brings 10 years of sales, marketing and consulting experience to KANA. As director of product marketing, Chopra is responsible for driving effective marketing programs and creating strategies for advancing the KANA eService Solutions in the marketplace. Prior to joining KANA in 2003, he held instrumental positions at leading companies such as Oracle, Commerce One and Ernst & Young, LLP.
What a Robot Really Wants

Knowledge at the Point of Decision

By Gordon Taylor, Senior Technical Consultant, TOWER Software

A small, two-legged robot stands atop a glass-topped coffee table. On its two dimensional world, it has to contend with a potted plant, an old TV Guide and several coffee cups, along with the ever-present danger of plummeting over the side onto the carpet below. As the robot navigates its way around, a constant evaluation process occurs inside its software “brain.” First, information is collected through its sensors. Secondly, that information is analyzed, using a decision tree to determine the optimal course of action.

What do the adventures of this robot have to do with knowledge management? It’s more important than you might think. You see, right before the robot takes its next step—one this simple two-stage process is completed—the robot could be said to “know” something. It’s collected all available information and analyzed it. Knowledge is created through analysis of information.

So the effectiveness of our robot friend—or if you like, how “smart” it is—depends directly on two things: the accuracy and relevance of the information supplied; and the effectiveness of the evaluation process. Poor information, through faulty sensors or too few sensors, will result in an inaccurate picture being fed to the decision-making processes. Poor analysis will lead to bad decisions, regardless of the quality of information supplied.

Now I’m sure you saw this analogy coming, but face it—your enterprise is exactly the same. To create a smart enterprise, you need to have a stable, reliable information base and the analysis tools that allow you to create valuable knowledge—knowledge that fosters good decisions.

Information management has been refined over the years, to the point where most enterprise architects are including a central structured repository as part of their information architecture. ECM systems, built on solid data-storage solutions, are the platforms that facilitate these sound information management policies.

At the heart of these information systems, is metadata—data stored about the data you store. By monitoring, storing and indexing specific information about your business content, ECM vendors allow their customers the ability to easily find any piece of information, and its relevant business context, quickly and efficiently. These systems are built on information management policy and principles that have been around for a long time.

So, if your organization has a sound ECM policy and system in place, it’s not likely to fall off the coffee table because of poor quality information. The next generation of enterprise systems will focus on how to manage the analysis of that information base to support your decision-making process.

The DIKW Model is an information hierarchy that’s frequently cited when trying to address this problem. The model was originally recorded in a 1932 poem called The Rock from TS Eliot:

Where is the life we have lost in living?
Where is the wisdom we have lost in knowledge?
Where is the knowledge we have lost in information?

In the modern, slightly less poignant implementation of the DIKW model, we find four layers:

- Data:
- Information:
- Knowledge:
- Wisdom.

“Poor analysis will lead to bad decisions, regardless of the quality of information supplied.”

Nowadays, thanks to advances in data storage, the science of information management and the implementation of these systems in ECM products, the transition from data to information is largely a solved problem.

Getting from information to knowledge is much more difficult. Knowledge includes the “how” aspect of a problem. Returning to our robot, it’s the analysis of the information that tells it “how” to proceed.

Current efforts at solving this problem are varied, and you’ll probably recognize them as the more modern features provided by ECM vendors: collaboration, which allows people to discuss and share information in order to facilitate progress; and workflow, where prescriptive, best-practice knowledge as defined by a business process analyst provides “how” information. Content management tools, like blogs and wikis, all provide additional published content around a topic—more published analysis to help people decide which step to take next.

Tools like these are striving to bridge the conceptual void between information and knowledge. While the jury is still out on how effective they are, the challenge is considerable. The next time you need to evaluate a system for inclusion in your enterprise architecture, consider how well it bridges this gap. Think like a robot. Do I have the right information available? Will this system enable me to make better decisions? Without a careful approach to both aspects, you could end up on the carpet.

TOWER Software, a leading enterprise content management (ECM) provider to government and regulated industries, delivers award-winning information management solutions. Our product, TRIM Context 6, enables organizations to have compliant, secure and accurate information available to make confident business decisions. TRIM Context 6 won AIIM E-DOC Magazine 2006 Best of Show award for ECM suites.
Enterprise Search: The Foundation for Risk Management

By Laurent Simoneau, CEO, Coveo Solutions, Inc.

If you can’t find it, you can’t manage it. Unfortunately, many executives have discovered this too late in the game—when they face regulatory compliance penalties, legal troubles or a corporate scandal. Today, there is no question that managing security risks and complying with government regulations related to information security can be a daunting task for any organization. In fact, many enterprises now have a mandate to implement technologies and applications that ensure all employees, not just corporate management, can access appropriate information securely.

Having a comprehensive, highly secure enterprise search capability—one that fills the gap between specialized search systems and Web-focused search tools—can be a key business asset, and is essential to effective knowledge management for corporations and government entities. When enterprise search has a strong emphasis on knowledge management, intellectual property, e-discovery and compliance, it becomes the foundation for comprehensive risk management.

Whether it is to meet regulatory compliance requirements from Sarbanes-Oxley (SOX), to the Health Insurance Portability and Accountability Act (HIPAA) or to respond faster to e-discovery requests, companies of all sizes today need to be able to search through literally gigabytes and terabytes of stored data wherever it resides. In other words, companies have to find what they are looking for, every time, all the time— their future may depend on it.

The Role of Search in Compliance

For the California Conservation and Liquidation Office (CLO), a division of the California Department of Insurance responsible for regulating all insurance companies operating in California, the compliance task requires managing insolvency, asset distribution, claimant lawsuits and disbursements. The office also has to routinely determine if an insurance company can be rehabilitated and continue its operations. This requires the CLO to keep extensive forms, policies and change control documentation on an intranet available to both employees and state auditors. And to ensure compliance, it must keep all processes and change controls current, and make sure each step in any procedure is well-documented and implemented without deviation.

According to CLO’s Mohammed Mojabi, “As a highly regulated entity, we are in a constant state of audit so we need a search engine that is highly secure, highly accurate, yet easy to use for both users and auditors.”

With CES, users and auditors quickly find current forms and documented procedures, reducing the time it takes to find and act on information by up to 35%. In addition to the time and money saved, CES minimizes compliance violations and results in far more efficient audits.

The Role of Search in Litigation

In today’s litigious world, companies must respond not just quickly, but quickly enough to litigation requests for electronic evidence...or face heavy court-imposed sanctions. The problem? e-discovery is literally like finding a needle in a haystack.

ProSearch Strategies, a discovery, analytics and workplace tools company focused on the legal market in areas of litigation, due diligence and compliance, faces these issues every day. During its due-diligence process, it follows a highly accurate automated process to minimize the manual review of documents—which is time-consuming, error-prone and expensive. For a recent case, ProSearch had to search through some 27 terabytes of unstructured data and needed an accurate technology solution that would enable it to hone in on the right data, quickly. ProSearch Strategies settled on a combination of applications including Microsoft SharePoint, SQL Business Analytics and Coveo Enterprise Search (CES) for back-end processing on the data. CES is the foundation application that allows ProSearch Strategies to access volumes of information in a manageable, structured format.

“We’re using software-as-a-service and SharePoint as the secure portal structure on the front-end, and SQL business analytics and Coveo for back-end processing on the data. This solution has enabled analysts to consolidate and eliminate time-consuming analytic tasks, decreasing processing time by 30%, and allowed researcher-reviewers to make more document decisions per hour, while maintaining the ease of use for casual searchers,” says Trevor Allen, CIO of ProSearch Strategies.

Find information. Understand Information. Act—much faster. Based on industry standard ASP and .NET technologies, and winner of the 2006 Microsoft Partner Regional Winning Customer Award, Coveo Enterprise Search delivers the best value in the marketplace with out-of-the-box document level security, unparalleled accuracy, consumer style ease of use, and an implementation cycle of less than 24 hours. Whether it’s to meet regulatory compliance, improve customer response, protect intellectual property or improve organizational efficiencies up to 35%, Coveo Enterprise Search enables organizations to find, understand and take action on critical information located anywhere in the enterprise.
Realizing Measurable ROI with Multi-Language Content Management

By Jon Parsons, XyEnterprise

What is content management worth?

That can be a philosophical question. If you ask a content author, you’ll hear about the way content creation and review is made easier and more efficient. If you ask a project manager, you’ll learn that content management makes it possible to identify bottlenecks and monitor the progress of component pieces in an information set. If you ask a production person, you’ll discover that delivering to different data formats is now a more automated and reliable process. Ask an IT person and they’ll talk about integration with existing enterprise infrastructure. While all of these answers are true, none of them really assigns a dollar value to content management. Finding a concrete way to measure ROI for content management can be a challenge.

One place to look for a metric to measure the value of content management is in the rising costs associated with delivering product-related content in many different languages. The global economy is here. Products that were once marketed and sold in areas that required only one or two languages are now being offered worldwide. Suddenly the Pacific Rim, Eastern Europe, India, China, and Brazil are growing markets for many kinds of products and services. To reach them and sell to them successfully, content such as marketing and sales information, product documentation, services manuals, warranty policies and many other kinds of content must be available in each of the many languages spoken there. Translation is the key. And time-to-market is critical. The required content is already in one language. It just needs to be put into the languages needed for your new markets.

The Rising Cost of Translation

But translation is expensive. Whether you translate in-house or use a translation service, whether the process is aided by a translation memory system or not, no matter how it gets done, translation has an assigned cost that can be readily determined for your business. And, typically, translation costs are rising—not because the per-page rates are increasing, but because more and more content needs to be translated. Product lifecycles grow shorter. Many more variations of products are introduced. New products are added to those already there. The increasing amount of content required to support these products means that more must be translated and, inevitably, translation expenses rise.

So, in calculating the ROI on introducing content management, you can get an objective metric to start with—the amount you currently spend on translation. This starting point is good for either projecting what you will save over a given period or analyzing the results of a pilot project designed to show what you can achieve. But still, that’s only half the story. How do you measure what your savings are, or what they could be?

Translate Once, Translate Right

The answer lies in the ability of multi-language content management to control what gets translated. If you have content that evolves from one product release to the next, chances are not everything in the content related to that product changes. Some things do. Maybe many things do. But not everything.

The key to controlling translation costs is to translate only those things that have changed. Your savings come from not translating those things that haven’t. Factor in the savings in time and resources and that number grows exponentially.

But in order for this approach to work, your content must be managed in granular units, sometimes called “minimum reusable units.” An XML-based approach like the Darwin Information Typing Architecture (DITA) calls them “topics.” Once your content is authored in such units, a content management system can determine when one of those units changes. When that content goes through the review and approval cycle and becomes ready for the next edition, a tool in the content management system can identify all the related units in other languages that must now be retranslated to capture the new or changed content. If the base unit is brand new, the system can also determine the need for creating new objects in the target languages. Such a content management system can then create a project and send the new content along with the older translated version off to the translation process.

When the translated units return, they are put in the repository and managed with all the others so that when an information set in the target language is published, it is easy to verify that all units are at the current version and in sync with the source language.

This approach to controlling translation costs is demonstrable and in use today. How does this approach provide a metric for measuring ROI? By using it, you can readily determine the percentage of content that is reused and not translated. The translation costs that would have been expended provide one good measure of your ROI. If that alone justifies your investment cost, then the other less readily measurable benefits of multi-language content management come along for free.

XyEnterprise is a leader in content management and multi-channel delivery, providing solutions to industry leaders in technology, financial services, publishing, manufacturing, government and aerospace/defense. The company continues to innovate to deliver products that support an evolving marketplace—including distributed workflows, multi-language content management, interactive electronic delivery and standards such as DITA and S1000D. Their solutions help workgroups of all sizes simplify and expedite the automated creation, management, delivery and reuse of content across the enterprise. info@xyenterprise.com

Jon Parsons has more than 20 years of experience automating the creation, management and delivery of content in multiple forms. Currently he works in product marketing at XyEnterprise. Prior to that, he was a writer, editor, tools developer and publishing consultant for a large computer manufacturer. Long an advocate of generic mark-up and an enthusiast for XML, he has served on the board of directors of OASIS, the Organization for the Advancement of Structured Information Standards, and is a frequent speaker at industry events.

KMWorld November/December 2006  S19
Hosted XML Content Management: Is It Right for You?

By Dan Dube, Managing Director, US Operations, DocZone.com

Publishing organizations have long recognized the value of migrating content to XML to attain the benefits of content reuse, reduced localization costs and single-source publishing. But, many of these organizations have never been able to justify the high cost and long implementation cycles required to install an in-house XML content management system (CMS).

Recently, a new alternative has emerged—the “Software as a Service (SaaS)” model, which offers a hosted XML content management environment on a subscription basis. According to research firm InfoTrends, more than 40% of their survey respondents would either “prefer” or “definitely consider” a hosted content management solution.

Software as a Service: What is It; Why Should I Care?

The SaaS business model is essentially designed to offer a full-featured solution in a hosted environment. The software application sits in a centralized, secure data center and is served up to end users completely via a browser. Rather than buying and implementing an expensive in-house solution, the customer pays a subscription fee to use licenses on the system. The vendor has to perform to the specifications of a service level agreement (SLA), or there are typically financial penalties to pay. SaaS is gaining acceptance as an alternative business model, led by the popularity of applications like WebEx and Salesforce.com.

Any organization can benefit from an SaaS business model, regardless of its size:

◆ Small-to-midsize business: For small-to-midsize businesses, SaaS allows access to software that might otherwise be too costly or complex to implement or support.
◆ Enterprise: For larger organizations, SaaS allows departments to avoid having to make large capital expenditures and having to pay for internal support costs. Large corporate environments typically turn to SaaS to support short-term projects, software that will only be used occasionally or by a small number of employees and for applications that need to be available outside of a firewall to partners, contractors, suppliers or customers.

Traditional CMS vendors typically charge most, or all, of the purchase price at the time a contract is signed, before the system is even installed. Usually, the customer is responsible for the system deployment (often working with a consulting/integration firm). The vendor charges 18% to 20% annually for software support, and is not accountable for implementation failure, even if the system is never actually used in production.

By contrast, an SaaS vendor is responsible for configuring the environment and delivering a turnkey application. License fees to an SaaS vendor do not start until the system is production-ready, and there are financial penalties for failure to meet the metrics in the service level agreement.

Many “traditional” CMS vendors are considering (or announcing) that they will now support a hosted model as an alternative delivery mechanism. Most of these companies will struggle, because they will now be held more accountable for a successful production implementation and will have a difficult time adjusting to having to wait for payment. It will also be very hard for these companies to give up their ongoing profitable maintenance revenue. (For example, 45% of Oracle’s revenue comes from maintenance.)

Implementing or upgrading a content management environment is a significant and risky undertaking, and there are many options available for consideration. But if you properly define your business needs, stick with solutions that conform to standards, start with a small pilot project and look for rapid ROI models (such as SaaS), your chances for success will increase dramatically. Choose wisely—the rewards are well worth it.

Example of an SaaS CMS: DocZone.com

DocZone.com provides the first commercially available XML content management platform available exclusively with the SaaS “on-demand” business model. Our customer base spans many industries, from automotive to hardware/software manufacturers to healthcare solution providers to utilities. Some examples include:

◆ A European automotive company is using DocZone to facilitate the creation, localization and automated publishing of glove box manuals in up to 30 languages, including bidirectional languages such as Arabic;
◆ A global healthcare company is implementing DocZone to manage the editorial, localization and single-source publishing of technical manuals, Web-based training materials and HTML help from the same set of source content; and
◆ A localization provider is using the DocZone platform to facilitate its translation and content optimization services to its end-user clients, allowing them to pass on significant savings for translation and desktop publishing of multilingual content and making it a more competitive player in the localization industry.

DocZone.com is a privately held company headquartered in Heemstede, the Netherlands, with a wholly owned US subsidiary headquartered in Bedford, New Hampshire. DocZone.com has direct employees in the Netherlands and US, as well as a close network of development and integration partners. DocZone.com is a KMWorld Trend-Setting Product of 2006.
Who is That “he?”

Using Pronouns and Anaphors in Text Extraction

By Gregory F. Roberts, International Sales & Marketing, AeroText, Lockheed Martin Integrated Systems and Solutions

Text extraction is a powerful tool to find and categorize elements in unstructured documents. These elements, or entities, are connected together to form the relationships, facts and events in a document. Oftentimes, the surface forms of an entity are not sufficient to capture and glean all of the necessary information for further analysis. Text extraction needs to be able to capture and link three underlying pieces of information to rightly categorize elements in a document: pronouns, anaphors and inferable attributes. This article will first describe how leveraging pronominial information improves text extraction. Second, I’ll explain how anaphors are useful to text extraction. Finally, this article will describe how inferable attributes are used to further enhance text extraction.

Pronouns

Very often in a document, a relationship occurs that cannot be determined solely from its surface structure. Pronouns, such as he, she and it, are only valuable if they can be related back to the entities which they refer. Consider the following example.

he was the operational chief of the organization

This phrase has little value for text extraction unless one can determine who the he refers to. Once a text extraction tool can determine pronominal reference, relationships between entities become more apparent. Now consider the previous example when one knows the pronominal referent.

he[Ali Ghufron] was the operational chief of the organization

A relationship that Ali Ghufron is the operational chief of the Jemaah Islamiyah is now apparent. But is there any more information we can glean from this example?

Anaphors

Anaphors are referents to some other entity within a document. Another type of anaphora that is not discussed in this article is exophora. Exophora is when the referent requires real-world knowledge and lies outside of the document. While pronouns mentioned above are also anaphors, referents are typically other entities or noun phrases. Look again at our previous example.

he[Ali Ghufron] was the operational chief of the organization

An anaphor reference capability can enhance a text extraction tool. Relating referents to anaphors provides additional information when discovering relationships, facts and events. Now consider the previous example when one knows the referent to the organization.

“An anaphor reference capability can enhance a text extraction tool.”

he[Ali Ghufron, GENDER=Male] was the operational chief of the organization [Jemaah Islamiyah]

A relationship that Ali Ghufron is the operational chief of the Jemaah Islamiyah is now apparent, even though the relationship was never explicitly stated in our document.

Inferable Attributes

Inferable attributes are those things about an entity that can be determined from other bits of the document or are implicit in the entity itself. For example, the gender of a name like John Smith can be inferred to be male simply because one knows that the name John tends to refer to males. However, some names are ambiguous when it comes to their attributes. From our example, the gender of the name Ali is ambiguous. The name is used for both males and females. One could merge the attributes of all occurrences of an entity together to make a composite entity. We learned that the pronoun he is linked together with Ali Ghufron in our previous example. Since he is a pronoun that refers to males, one can assume that the entity Ali Ghufron is also male. Reconsider our previous example.

he[Ali Ghufron] was the operational chief of the organization [Jemaah Islamiyah]

We now know that the male Ali Ghufron is the operational chief of the Jemaah Islamiyah, even though the relationship was never explicitly stated in our document.

Structured Language

In this article, I showed how leveraging pronominal information can provide more meaningful results to a text extraction tool. I also demonstrated how anaphors add value to text extraction by making relationships become more apparent. Finally, I described how a text extraction tool could use inferable attributes, either implied from the entity itself or by merging the attributes of all the various occurrences of the entity, to further enhance text extraction. These are just a few examples of how the underlying structures of natural language can be used by text extraction tools to enhance the value of extracted information for real-world applications.

Headquartered in Herndon, VA, Lockheed Martin’s Integrated Systems & Solutions (IS&S) was formed in June 2003, in response to the increasing demand for solutions that promise a comprehensive, real-time information picture for faster, better informed decisions. Developed with more than 20 years of Lockheed Martin experience, AeroText is a high-performance data extraction engine and development environment that worldwide companies and governments use to find and correlate relevant information in text documents.
Beyond First Call Resolution

Diagnostic and Measurement Practices for KM

By Andrew Cohen, Ph.D., Sr. Director of Business Consulting, KNOVA Software

Knowledge management (KM) initiatives are one way of improving technical support organizations, driving support margins through efficiency and increased customer loyalty. However, it is one thing to implement different practices in KM in your organization and another to know that the practices are contributing to your organization’s bottom line. In what follows, we outline best diagnostic and measurement practices for judging the effects of new KM practices.

Increasing “first contact resolution” (FCR) is often one of the first goals specified in a project. Organizations often believe a significant reason their satisfaction is low and costs are high is that calls or cases are not solved on the first contact. They regularly point to long elapsed times for cases, poor satisfaction scores and lowered support margins as indicators that they need to improve the expertise of their front-line support personnel and the explicit knowledge that supports their efforts. As we shall see below, this is only partly true. Pursuing this path alone may not fully address the FCR problem and the corresponding cost/satisfaction issue.

Following common best practices, let’s examine the situation:

1. Map and model current process: For complex technical support centers, there are multiple interactions or activities, often including three or more touch points with the customer, before any troubleshooting starts. Support engineers may require logs, configuration and sample files just to start their assessments. Adding expensive experts to your front line will only exacerbate the problem. The experts will have to go through the same set of iterations and activities with customers before they get started.

2. Establish metrics: Measuring first call resolution alone does not give you the whole picture. In this case, you need to measure the number and content of the iterations. How much time and over how many iterations does it take an engineer to get the appropriate information from the customer? What is the average elapsed time that cases are open, and what are the cut-off points for customer delight and customer pain?

3. Perform a diagnostic: We often find that the multiple iterations required before troubleshooting can get started do not need to be handled by an expert. In the vast majority of cases, there are technology solutions. In our experience, we can automate the collections of the knowledge required to initiate troubleshooting. Self-service wizards can walk the customer or front-line agent through the critical information required to initiate troubleshooting.

4. Set clear goals: Monitor the underlying processes. Monitor your customers’ and agents’ quantity and success with the wizards, the elapsed time cases are open and how often the users (customers or agents) get the correct log info to the engineer before they start the troubleshooting process.

5. Develop your future state: Only for some of the problems can wizards collect log and configuration information. Specifying the set of applicable problems for which this is possible is a pre-requisite to quantifying the benefit.

6. Quantify the benefit: Develop word equations to specify the savings. For example, the potential savings from collecting logs automatically is (% of time engineers spend collecting sys files) X (% of time sys files are needed) X (% of time the collection could be completed in a wizard) X (% of time cases are submitted online). The result of this is an efficiency improvement for your engineers. Similar word equations assess the reduction in elapsed time cases are open.

7. Run a continuous improvement process: Using technology to solve problems requires ongoing effort. Make sure you are assessing which problems are likely to apply to this process. Set 30- to 90- and 120-day assessment periods and add/modify wizards as necessary.

You can see that low FCR is a proxy for a host of underlying processes. Measuring your process today and theorizing how your new processes will help you close gaps that underlie problems can help remedy core problems. Continuing to improve and benchmark processes will help you to know how changes improve your organization. Over time incrementally improving and benchmarking improvements will help you reach your customer satisfaction and efficiency goals.

Andrew Cohen, Ph.D. is the senior director of business consulting at KNOVA Software. His team is responsible for business strategy and process re-engineering in pre-sales, implementation and post-implementation services. Prior to joining KNOVA, Cohen ran the knowledge management and distributed learning research team at IBM Research’s Cambridge lab. Cohen received a Ph.D. in cognition and learning from the University of Toronto and has a Masters of Science and BA in physics.

KNOVA Software is a leading provider of intelligent customer experience solutions that maximize the value of every interaction throughout the customer lifecycle. Built on an adaptive search and knowledge management platform, KNOVA’s suite of applications helps companies increase revenues, reduce service costs and improve customer satisfaction. Industry leaders including AOL, Ford, HP, Novell, Reuters, McAfee and H&R Block rely on KNOVA’s award-winning applications to power an intelligent customer experience on their Web sites, and within their contact centers. For more information, visit www.knova.com.
The Hidden Costs of Product Information Publishing

By Chip Gettinger, Vice President, Customer Service and Support, Astoria Software

Manufacturing enterprises have placed a sustained focus on information management solutions to support and augment the design, development and production process. The associated product and manufacturing information that parallels the production process is closely managed, tracked and reviewed to optimize the overall product lifecycle opportunity.

Increasingly, many organizations are tracking and measuring the product information chain beyond the direct manufacturing process. The findings? This extended chain is much longer, broader and costlier, with a decentralized network of product information creators and consumers. The critical nature of this chain is evident as most content flows directly to the creation of customer-facing product information, from sales and marketing content, to product documentation and operations manuals, and to customer service and support information. Some estimate the current costs of content creation, management and delivery throughout the organization accounts for up to 6% of the costs of goods sold for any one product.

Volume, Velocity and Variability

As companies add to and customize product lines, the associated content costs grow exponentially. The volume of content rises as product lines are extended, the velocity of content accelerates as product lifecycles shorten and the variability in information grows as products are customized and internationalized. The strain on the organization’s product information processes is evident as squeezed product documentation deadlines and missed launch dates, information inaccuracies that trigger legal liability and increased redundancy costs as content is updated and transformed for multiple languages and delivery formats.

Enterprise customers now identify and measure where value and costs are created or diminished throughout the product information lifecycle. From original content creation, to review and translation cycles, to document assembly and delivery, there are four specific areas identified within organizations where content value and cost is likely to be measured: re-use of content; collaboration and review workflow cycles; language translation processes; and publishing.

What is clear is that the more an organization can eliminate a redundant, standalone approach to product information processes, the better the cost-efficiencies across the entire content lifecycle. For many, the old model of writing and locking valuable content in static documents stored on file-servers in a business group is crippling in an age when agility to respond is critical to success.

A Standards-based Approach

Astoria Software has helped leading organizations realize efficiency and cost-savings in how they create, manage and deliver their product information. Most start with the adoption of a single-source information model, relying on a centrally managed content repository. This centralization of content in an XML repository delivers flexibility as an organization creates, tracks, updates, translates and publishes content and manages its re-use over time. Besides automation tools, organizations are adopting information standards such as the OASIS Darwin Information Typing Architecture (DITA). This standard supports a single-source information model with the creation and management of content “topics” that are easily re-used. Content topics, along with metadata and product content attributes are managed independently, but maintain relationships with each other. DITA’s granular topic-based approach, in concert with an XML content repository, delivers inherent flexibility in the creation of documents, and supports efficiencies throughout the author-to-publish process.

As demand for new products continues to put pressure on content groups, the need for more cost-effective management of the information lifecycle becomes critical to business efficiency. Solutions such as Astoria’s have helped Fortune 500 companies streamline their product information processes and reduce costs, from $3 million to $300,000 in documentation translation costs alone for one Fortune 500 manufacturer. It starts with an organization’s capability to re-use content, to effectively collaborate, and to automate multi-language, multi-format publishing. Finally, it means adopting a single-source information model and content management solutions that can effectively manage the entire process in a measurable and cost-controlled way.

Chip Gettinger is a long-time executive in the information publishing industry and is a regular speaker at content management and publishing industry conferences. As vice president of services and support for Astoria Software, he advises customers on best practices in product information management. Gettinger is also involved with OASIS, the XML standards board, and is leading the first DITA specialization committee for medical manufacturing. Astoria, based in San Mateo, CA, and founded in 1994, is a leading provider of XML content management solutions for the dynamic publishing of product documentation and content.

To learn more about how Astoria has helped leading organizations including Siemens Medical Solutions, Texas Instruments, NCR Teradata and more, visit www.astoriasoftware.com/kmworld, or via email at info@astoriasoftware.com.

Technical Document Publishing On-Demand

A manufacturer’s technical publication departments are most affected by the velocity of new product innovation, and the volume of demand for new content delivered in more formats and languages.

To address this need, Astoria Software introduces a new end-to-end XML content management solution now available as an on-demand offering. This best-in-class solution delivers an XML authoring tool, an XML content management repository, workflow, a built-in DITA Workbench for simple transition to the DITA standard and a composition engine to output to any delivery format—print or digital. This entire solution is accessed from the Web via a third-party hosting provider, eliminating the need for any hardware or services investment, and supported and serviced by Astoria. For more information, visit Astoriasoftware.com/OnDemand.
For more information on the companies who contributed to this white paper, visit their websites or contact them directly:

**Fast**

Fast Search & Transfer, Inc.
117 Kendrick Street, Suite 100
Needham MA 02494
PH: 888.871.3839
FAX: 781.304.2410
Contact: info@fastsearch.com
Web: www.fastsearch.com

**InQuira**

InQuira Inc.
851 Traeger Avenue, Suite 125
San Bruno CA 94066
PH: 650.246.5000
FAX: 650.246.5036
Contact: www.inquira.com/contact.asp
Web: www.inquira.com

**KANA**

KANA Software
181 Constitution Drive
Menlo Park CA 94025
PH: 800.737.8738
Contact: sales@kana.com
Web: www.kana.com

**KNOVA**

KNOVA Software
10201 Torre Avenue, Suite 350
Cupertino CA 95014
PH: 800.572.5748
Contact: info@knova.com
Web: www.KNOVA.com

**Lockheed Martin**

Lockheed Martin Corporation
Integrated Systems & Solutions
13560 Dulles Technology Dr.
Herndon VA 20171
PH: 703.466.1268
Contact: AeroText.sales@lmco.com
Web: www.aerotext.com

**Noetix Corp.**

Noetix Corp.
5010 148th Avenue NE, Suite 100
Redmond WA 98052
PH: 866.4NOETIX
FAX: 425.372.2942
Contact: sales@noetix.com
Web: www.noetix.com

**SchemaLogic Inc.**

SchemaLogic Inc.
620 Kirkland Way, Suite 100
Kirkland WA 98033
PH: 425.885.9695
FAX: 425.883.0117
Contact: sales@schemalogic.com
Web: www.schemalogic.com

**TOWER Software**

TOWER Software
12012 Sunset Hills Road
Two Discovery Square, Suite 510
Reston VA 20190
PH: 800.255.9914 or 703.476.4203
FAX: 703.437.9006
Contact: info@towersoft.com
Web: www.towersoft.com

**XYEnterprise**

XYEnterprise
101 Edgewater Drive
Wakefield MA 01880
PH: 781.756.4400
FAX: 781.756.4330
Contact: info@xyenterprise.com
Web: www.xyenterprise.com

**DocZone**

DocZone.com
21 McAfee Farm Road
Bedford NH 03110
PH: 603.488.5008
FAX: 603.488.5009
Contact: info@doczone.com
Web: www.doczone.com

**BEA Systems, Inc.**

BEA Systems, Inc.
500 Sansome Street
San Francisco CA 95111
PH: 800.817.4BEA (US toll free)
Contact: info@bea.com
Web: www.bea.com

**Coveo**

Coveo Solutions Inc.
120 Hawthorne Avenue, Suite 100
Palo Alto CA 94301
PH: 800.635.5476
FAX: 650.475.8024
Contact: info@coveo.com
Web: www.coveo.com

**Endeca Technologies, Inc.**

Endeca Technologies, Inc.
55 Cambridge Parkway
Cambridge MA 02142
PH: 617.577.7999
FAX: 617.577.7766
Contact: sales@endeca.com
Web: www.endeca.com

**KANOA Software**

KANOA Software
10201 Torre Avenue, Suite 350
Cupertino CA 95014
PH: 800.572.5748
Contact: info@knova.com
Web: www.KNOVA.com

**Lockheed Martin**

Lockheed Martin Corporation
Integrated Systems & Solutions
13560 Dulles Technology Dr.
Herndon VA 20171
PH: 703.466.1268
Contact: AeroText.sales@lmco.com
Web: www.aerotext.com

**Noetix Corp.**

Noetix Corp.
5010 148th Avenue NE, Suite 100
Redmond WA 98052
PH: 866.4NOETIX
FAX: 425.372.2942
Contact: sales@noetix.com
Web: www.noetix.com

**SchemaLogic Inc.**

SchemaLogic Inc.
620 Kirkland Way, Suite 100
Kirkland WA 98033
PH: 425.885.9695
FAX: 425.883.0117
Contact: sales@schemalogic.com
Web: www.schemalogic.com

**TOWER Software**

TOWER Software
12012 Sunset Hills Road
Two Discovery Square, Suite 510
Reston VA 20190
PH: 800.255.9914 or 703.476.4203
FAX: 703.437.9006
Contact: info@towersoft.com
Web: www.towersoft.com

**XYEnterprise**

XYEnterprise
101 Edgewater Drive
Wakefield MA 01880
PH: 781.756.4400
FAX: 781.756.4330
Contact: info@xyenterprise.com
Web: www.xyenterprise.com

**DocZone**

DocZone.com
21 McAfee Farm Road
Bedford NH 03110
PH: 603.488.5008
FAX: 603.488.5009
Contact: info@doczone.com
Web: www.doczone.com