

Knowledge Management Technology Solutions

Oct 2000

Introduction

Successful companies have made their mark because they not only had the knowledge but were also able to harness it in order to make critical business decisions. In this era where information is prolific and overwhelming, the winners are going to be those that are able to store, process, and formulate knowledge from the chaos, operate upon it, and share it within the company. According to the Mercer Marketplace 2000 Survey, 80% of all respondents believe transforming information into knowledge is the second most important source of competitive advantage, second only to establishing and maintaining customer relationships.

Profitability is probably the strongest driver compelling companies to adopt knowledge management practices. Other drivers are globalization, time-to-market pressures, information overload, employee attrition, and stronger competition. Those companies that are using knowledge management practices today have identified the following key benefits:

- ?? Improved decision-making,
- ?? Increased responsiveness,
- ?? Improved efficiency of people and operations,
- ?? Improved innovation,
- ?? Improved products/services, and
- ?? Greater flexibility.

Indeed, companies investing in knowledge management, such as IrisLogic, are getting significant advantages from implementing knowledge management solutions. In a survey conducted in 1997 by the Journal of Knowledge Management in conjunction with The Benchmark Exchange and the Best Practice Club, respondents recognized the costs associated with knowledge-based errors and mistakes. Practitioners dealing with day-to-day operating problems were more aware of the costly mistakes due to "knowledge problems" than their top management teams.

Some 97% of the respondents reported critical business processes which were "owned" by only one or two people, and 87% of the surveyed group stated that they were personally aware of costly mistakes because employees lacked sufficient knowledge or expertise, or knowledge was not available when and where needed.

Why Knowledge Management

At a time when firms need to "know what they know" and must use that knowledge effectively, the size and geographic dispersion of many of them make it especially difficult to locate existing knowledge and get it to where it is needed. In a small, localized company, a manager probably knows who has experience in a particular aspect of the business and can walk across the hall and talk to them. Studies have shown that the maximum size of an organization in which people know one another

well enough to have a reliable grasp of collective organizational knowledge is two hundred to three hundred people. In a global enterprise, with offices scattered about the planet, how does one access the in-house knowledge that they need? How does a group about to launch a new product know about the lessons learned from another group who have just done the same thing? The mere existence of knowledge somewhere in the organization is of little benefit; it becomes a corporate asset only if it is accessible.

Studies have also shown that managers get two thirds of their information and knowledge from face-to-face meetings and only one third from documents. Collaboration between groups involving knowledge exchange allows for improvements on processes, prevents the wasteful "reinventing the wheel" scenario, and heightens innovation and productivity. Collaboration and teamwork revolve primarily around people forming networks to solve issues based on discussions where knowledge is shared from one to another. These discussions traditionally happen on an informal, ad-hoc basis. The problem with such impromptu discussions is that they are localized and fragmented. The move to "virtual offices" and telecommuting compounds the problem as fewer and fewer of these informal face-to-face conversations occur.

Companies need to encourage groups of people that share common work practices, interests, and aims to network together and share their insights. This can be done through e-mail, newsletters, or meetings. By establishing Competence Centers and Communities of Practice, others in the organization will have a focal point to which to turn with questions and issues in that area of expertise. In support of these collaborative organizations, tools such as IBM DB2 Knowledge X and Intelligent Mining for Text can be used to document, categorize, and deliver relevant subject material as well as a map to show people in the organization where to go when they need specific expertise.

Formalizing networking groups and formulating knowledge maps in and of themselves, however, is not sufficient. Companies need to foster a knowledge culture within their organization. Employees need to be rewarded for sharing their knowledge with others. Respect and trust in each other's knowledge is paramount to successful knowledge transfer and application. Open-minded discussions and a willingness to try new methods need to be encouraged--by their very nature, people resist any innovation that may require them to abandon their signature skills in favor of new ones.

Knowledge is a corporate asset, and an infrastructure is required to store it, maintain it, archive it, update it, discard it, and present it to the right people in the right format at the right time. Most companies have tapped into the information age with databases, data warehouses, and Internet search engines, and are primed to add the extension technologies to bring them into the knowledge age. With a few modern methods and tools and a shift in culture, an infrastructure can be created to manage and preserve Intellectual Capital.

The Value Chain: Data-Information-Knowledge

If knowledge is that valuable, what is it, how is it generated, and where can I get some? We all have access to data--tons of data. Too much data. Data is the raw

material we get from sources. When data is a set of discrete facts about business transactions, stored, for instance, in a database as a table (structure) with fields of numbers and literals, it is called "structured data". But documents and Web pages are also sources of data, delivered as a stream of bits that represent words and sentences of text in a certain language. Industry analysts estimate that data that has no external structure (internally there are paragraphs and words, but we have to find them) represent 80% of an enterprise information compared to 20% from structured data. This 80% comprises data from different sources, such as text, image, video, and audio. Text is, however, the most predominant variety of unstructured data.

Data in itself has very little meaning. When a customer goes to restaurant for dinner, this transaction can partially be described by data: when the customer ate there, what they ordered, how many courses were served, how much was paid, and how it was paid (cash, credit, check, debit card). This data tells nothing about why the customer went to that particular restaurant, cannot predict how likely it is that they will be back, whether the meal was good or not, nor whether the restaurant had good service. Data describes only a part of what has happened; it provides no judgment or interpretation and no sustainable basis of action. It has no inherent relevance or purpose. But data is the essential raw material for the creation of information.

Information has been defined by Peter Drucker as "data endowed with relevance and purpose." Information is usually described in a communications schema involving messages between a source and a receiver. These messages are coherently packaged data. Information is most valuable when it makes a difference for the receiver to know about it.

There is a value chain from data to information, and then to knowledge. Data is the result of a transaction, for instance, "5000 dollars". It does not have much relevance. Information is data coherently packaged to deliver a message. It makes a difference to receive it. For instance, "The monthly salary of John Wright is 5000 dollars" provides more meaning by packaging the value of the salary with the name of the employee.

The relevance of information for decision making and action depends on the context in which it is used. In this example, the information about John Wright's salary increase has a different relevance when he asking for a loan than when he estimating his taxes.

Textual data is usually semantically richer than numeric data because in text we rarely deal with individual words but with complete text documents that have meaning per se. For that reason, to access structured data, we talk about "data access", while to access a text document, we talk about "information retrieval", because each text document provides information. This information may not be relevant to the task at hand however. If we want to retrieve from the Internet documents referring to picture framing and we are not careful how we perform the search, we will end up with thousands of documents describing everything from cameras to the construction of houses. And even if we know how to establish the condition of the search in order to look for those words together, we will still end up with hundreds of documents related to picture framing, from which we have to sift those that are relevant to our task at hand. The enormous amount of information the user receives and is forced to digest has given birth to a new term: "infoglut".

Business users are "drowning in information", while starving for the insights that will allow them to make better decisions.

Well-managed information that is properly cataloged and structured, available and accessible by the right people at the right time becomes knowledge. Just as more tangible corporate assets like computer systems have a finite shelf life, so too does knowledge--it must be available at the right time to be able to act upon it. Although the concept of knowledge is hard to pin down, and is the object of many controversies, the reality is that we all recognize knowledge when we see it, be it in the head of an expert that tells us how to tackle a particular problem, or in a document that provides us with the required elements to solve that problem. A working definition of knowledge provided by the META Group is "a set of logical connections among pieces of information whose relationship is revealed through context and process familiarity" (10/97). We talk of "tacit knowledge"--experiential, latent in the head of the expert--and "explicit knowledge", when it can be transferred to a media that can be accessed by others.

Tacit knowledge is created in the heads of people. It is derived from exposure to experiences, data, and documents. It is hard to define and is usually shared through consultation, mentoring, and example. Retaining tacit knowledge means retaining the individual, which is not always possible.

Explicit knowledge is codifiable, and therefore can be made persistent. It can be generated from tacit knowledge by humans, for example, when the expert loan officer writes a screening procedure for loan risk assessment that is documented in procedure manuals. Alternatively, a knowledge engineer can elicit the business rules for the procedure from the expert and store them in the knowledge base of a rule-based system. However, human generation of explicit knowledge from tacit knowledge has always a complex, subtle and hard-to-pin-down process. The key ingredient for this exchange is the face-to-face sharing of knowledge or, in the automated world, through virtual environment tools like Lotus Notes™, which can become the focal point for tacit knowledge exchange. Future developments in simulation and virtual reality hold great promise as mediums for this type of knowledge transfer.

While face-to-face communication is the predominant method of sharing tacit knowledge today, technology that codifies and disseminates that knowledge, and indeed even creates knowledge, is increasingly being used in knowledge management infrastructures. Implicit knowledge can be discovered by computer processes known as machine learning, where the computer extracts knowledge from data and textual sources, using data mining and text mining tools. Text mining and data mining are used to extract previously unknown, comprehensible, and actionable knowledge from any source to make crucial business decisions. These sources include documents, images, e-mail, news feeds, and web pages.

Text mining tools play an important role in knowledge management infrastructures in that key knowledge can be extracted from textual documents, such as customer correspondence, on-line news services, e-mail, and Web pages. Predominant themes, patterns, associations, and trends are automatically uncovered in collections of documents and organized in such a way that users can grasp content quickly.

There are three key components to text mining: in-depth full text search engines, web solutions, and text analysis tools.

Full text search engines perform in-depth document analysis during indexing. They allow for sophisticated query enhancement and result preparation in order to supply high-quality information retrieval.

Web solutions can be used to build a global WWW search service or a centralized Intranet search service. They are designed and optimized to handle the large amounts of information. Therefore, document analysis and processing with a web search engine will be shallower than in-depth document search engines to provide for faster indexing and response to large volumes of queries. Web search engines enable users to create searchable indexes for the information on a Web site. With it, groups can make certain documents available on particular sites.

Text analysis tools are designed to be used in building applications that extract summaries and key concepts and organize documents according to user classification scheme. By using machine learning techniques the software trains itself to recognize similar or theme-related documents. Text analysis tools can capture multiword concepts in documents without using a domain dictionary. This feature coupled with intelligent clustering and categorization techniques enable knowledge discovery from textual sources. Often the first step to knowledge discovery is to extract key features from text, to act as "handles" in further processing. Examples of features are the language of the text, or company names and dates mentioned. After feature extraction, the next step may be to cluster documents into groups that are similar in content. These groups correspond to topics or themes discovered by the machine, not defined by the user. Categorization techniques enable the computer to assign documents to categories predefined by an established company-wide taxonomy. Categorization can also be used to dispatch requests, and forward news to subscribers.

Text mining in itself is a powerful technology that leverages the investments of an enterprise in knowledge management, by enabling automatic population of knowledge maps. Feature extraction techniques can help create a knowledge map from the contents of a document. Categorization can be used to populate knowledge maps with pointers to the right documents supporting a business process. Clustering enables the use of the discovered collection segments as knowledge objects of a map, and summarization is used for fast transfer of explicit knowledge. Text mining and knowledge mapping are therefore very important components of any knowledge management support infrastructure.

Metadata model generation is being heralded as the next evolution in knowledge management technology. The knowledge catalog is a categorized collection or repository of the company's intellectual assets, metadata provides the information about the assets, such as location, author, and the date of creation of the resource as well as the date that the resource was entered into the catalog. The metadata knowledge base requires a formal classification scheme, taxonomy, to coherently catalog knowledge. At an enterprise level, the amount of metadata can be huge, and require frequent accesses and updates. A company's knowledge catalog must be therefore built on a database platform that enables enterprise-wide scaling, protects and maintains the knowledge content, and allows for consistently fast access to the enterprise knowledge resources.

The key to successful knowledge management is personalization, or how to extract the knowledge that is pertinent to the user and translate it into a format that is

easily understood. Personalization is important not only for knowledge retrieval but also for navigation through all the various sources of data, information, and knowledge. Users are not necessarily aware of the existence of all the sources, nor do they know best how to navigate from one to another in a way that would be meaningful to their business needs. They need more than a super search engine-- they need a formal, disciplined mechanism to find their way through the hordes of information available. For example, consider all pieces of information pertaining to a company as items in a catalog. Certain types of information are relevant to one group within the company whereas a different set of information, perhaps overlapping the first, perhaps not, is relevant to another. If given an index of all items in the catalog, groups would have a hard time deciphering which items were relevant to them. What is needed is a mechanism that defines what items are relevant to which groups, creates searchable indexes and enables access to those items.

Knowledge maps coupled with metadata models are that mechanism. A knowledge map is a view of the knowledge catalog, allowing the visualization, manipulation, and navigation of the information and knowledge sources contained in the catalog. It uses the metadata layer to derive the connections from resource to resource. Conceptual indexes are abstracted from the complete catalog index, one for each group, enabling organized, disciplined navigation through information. Users are able to customize their knowledge maps, personalizing the information to their needs so that they can make immediate, effective business decisions.

Knowledge is an asset. Like any corporate asset, it can be lost or kindled. Paul Romer of Stanford University identifies knowledge as "the only unlimited resource, the one asset that grows with use." Knowledge can be likened to a living system, growing and changing as it interacts with the environment. Knowledge is created within a corporation through experiences and learning. It is created in the minds of people, as products are created in an enterprise. Ideas breed new ideas, and shared knowledge stays with the giver while it enriches the receiver. It is more valuable than many tangible assets and must be invested in and managed. This concept of successfully managing knowledge many would argue to be the Holy Grail for successful businesses.

The Value Chain: Data-Information-Knowledge

If knowledge is that valuable, what is it, how is it generated, and where can I get some? We all have access to data--tons of data. Too much data. Data is the raw material we get from sources. When data is a set of discrete facts about business transactions, stored, for instance, in a database as a table (structure) with fields of numbers and literals, it is called "structured data". But documents and Web pages are also sources of data, delivered as a stream of bits that represent words and sentences of text in a certain language. Industry analysts estimate that data that has no external structure (internally there are paragraphs and words, but we have to find them) represent 80% of an enterprise information compared to 20% from structured data. This 80% comprises data from different sources, such as text, image, video, and audio. Text is, however, the most predominant variety of unstructured data.

Data in itself has very little meaning. When a customer goes to restaurant for dinner, this transaction can partially be described by data: when the customer ate there,

what they ordered, how many courses were served, how much was paid, and how it was paid (cash, credit, check, debit card). This data tells nothing about why the customer went to that particular restaurant, cannot predict how likely it is that they will be back, whether the meal was good or not, nor whether the restaurant had good service. Data describes only a part of what has happened; it provides no judgment or interpretation and no sustainable basis of action. It has no inherent relevance or purpose. But data is the essential raw material for the creation of information.

Information has been defined by Peter Drucker as "data endowed with relevance and purpose." Information is usually described in a communications schema involving messages between a source and a receiver. These messages are coherently packaged data. Information is most valuable when it makes a difference for the receiver to know about it.

There is a value chain from data to information, and then to knowledge. Data is the result of a transaction, for instance, "5000 dollars". It does not have much relevance. Information is data coherently packaged to deliver a message. It makes a difference to receive it. For instance, "The monthly salary of John Wright is 5000 dollars" provides more meaning by packaging the value of the salary with the name of the employee.

The relevance of information for decision making and action depends on the context in which it is used. In this example, the information about John Wright's salary increase has a different relevance when he asking for a loan than when he estimating his taxes.

Textual data is usually semantically richer than numeric data because in text we rarely deal with individual words but with complete text documents that have meaning per se. For that reason, to access structured data, we talk about "data access", while to access a text document, we talk about "information retrieval", because each text document provides information. This information may not be relevant to the task at hand however. If we want to retrieve from the Internet documents referring to picture framing and we are not careful how we perform the search, we will end up with thousands of documents describing everything from cameras to the construction of houses. And even if we know how to establish the condition of the search in order to look for those words together, we will still end up with hundreds of documents related to picture framing, from which we have to sift those that are relevant to our task at hand. The enormous amount of information the user receives and is forced to digest has given birth to a new term: "infoglut". Business users are "drowning in information", while starving for the insights that will allow them to make better decisions.

Well-managed information that is properly cataloged and structured, available and accessible by the right people at the right time becomes knowledge. Just as more tangible corporate assets like computer systems have a finite shelf life, so too does knowledge--it must be available at the right time to be able to act upon it. Although the concept of knowledge is hard to pin down, and is the object of many controversies, the reality is that we all recognize knowledge when we see it, be it in the head of an expert that tells us how to tackle a particular problem, or in a document that provides us with the required elements to solve that problem. A working definition of knowledge provided by the META Group is "a set of logical connections among pieces of information whose relationship is revealed through

context and process familiarity" (10/97). We talk of "tacit knowledge"--experiential, latent in the head of the expert--and "explicit knowledge", when it can be transferred to a media that can be accessed by others.

Tacit knowledge is created in the heads of people. It is derived from exposure to experiences, data, and documents. It is hard to define and is usually shared through consultation, mentoring, and example. Retaining tacit knowledge means retaining the individual, which is not always possible.

Explicit knowledge is codifiable, and therefore can be made persistent. It can be generated from tacit knowledge by humans, for example, when the expert loan officer writes a screening procedure for loan risk assessment that is documented in procedure manuals. Alternatively, a knowledge engineer can elicit the business rules for the procedure from the expert and store them in the knowledge base of a rule-based system. However, human generation of explicit knowledge from tacit knowledge has always a complex, subtle and hard-to-pin-down process. The key ingredient for this exchange is the face-to-face sharing of knowledge or, in the automated world, through virtual environment tools like Lotus Notes™, which can become the focal point for tacit knowledge exchange. Future developments in simulation and virtual reality hold great promise as mediums for this type of knowledge transfer.

While face-to-face communication is the predominant method of sharing tacit knowledge today, technology that codifies and disseminates that knowledge, and indeed even creates knowledge, is increasingly being used in knowledge management infrastructures. Implicit knowledge can be discovered by computer processes, known as machine learning, where the computer extracts knowledge from data and textual sources, using data mining and text mining tools. Text mining and data mining are used to extract previously unknown, comprehensible, and actionable knowledge from any source to make crucial business decisions. These sources include documents, images, e-mail, news feeds, and web pages.

Text mining tools play an important role in knowledge management infrastructures in that key knowledge can be extracted from textual documents, such as customer correspondence, on-line news services, e-mail, and Web pages. Predominant themes, patterns, associations, and trends are automatically uncovered in collections of documents and organized in such a way that users can grasp content quickly.

There are three key components to text mining: in-depth full text search engines, web solutions, and text analysis tools.

Full text search engines perform in-depth document analysis during indexing. They allow for sophisticated query enhancement and result preparation in order to supply high-quality information retrieval.

Web solutions can be used to build a global WWW search service or a centralized Intranet search service. They are designed and optimized to handle the large amounts of information. Therefore, document analysis and processing with a web search engine will be shallower than in-depth document search engines to provide for faster indexing and response to large volumes of queries. Web search engines enable users to create searchable indexes for the information on a Web site. With it, groups can make certain documents available on particular sites.

Text analysis tools are designed to be used in building applications that extract summaries and key concepts and organize documents according to user classification scheme. By using machine learning techniques the software trains itself to recognize similar or theme-related documents. Text analysis tools can capture multiword concepts in documents without using a domain dictionary. This feature coupled with intelligent clustering and categorization techniques enable knowledge discovery from textual sources. Often the first step to knowledge discovery is to extract key features from text, to act as "handles" in further processing. Examples of features are the language of the text, or company names and dates mentioned. After feature extraction, the next step may be to cluster documents into groups that are similar in content. These groups correspond to topics or themes discovered by the machine, not defined by the user. Categorization techniques enable the computer to assign documents to categories predefined by an established company-wide taxonomy. Categorization can also be used to dispatch requests, and forward news to subscribers.

Text mining in itself is a powerful technology that leverages the investments of an enterprise in knowledge management, by enabling automatic population of knowledge maps. Feature extraction techniques can help create a knowledge map from the contents of a document. Categorization can be used to populate knowledge maps with pointers to the right documents supporting a business process. Clustering enables the use of the discovered collection segments as knowledge objects of a map, and summarization is used for fast transfer of explicit knowledge. Text mining and knowledge mapping are therefore very important components of any knowledge management support infrastructure.

Metadata model generation is being heralded as the next evolution in knowledge management technology. The knowledge catalog is a categorized collection or repository of the company's intellectual assets, metadata provides the information about the assets, such as location, author, and the date of creation of the resource as well as the date that the resource was entered into the catalog. The metadata knowledge base requires a formal classification scheme, taxonomy, to coherently catalog knowledge. At an enterprise level, the amount of metadata can be huge, and require frequent accesses and updates. A company's knowledge catalog must be therefore built on a database platform that enables enterprise-wide scaling, protects and maintains the knowledge content, and allows for consistently fast access to the enterprise knowledge resources.

The key to successful knowledge management is personalization, or how to extract the knowledge that is pertinent to the user and translate it into a format that is easily understood. Personalization is important not only for knowledge retrieval but also for navigation through all the various sources of data, information, and knowledge. Users are not necessarily aware of the existence of all the sources, nor do they know best how to navigate from one to another in a way that would be meaningful to their business needs. They need more than a super search engine-- they need a formal, disciplined mechanism to find their way through the hordes of information available. For example, consider all pieces of information pertaining to a company as items in a catalog. Certain types of information are relevant to one group within the company whereas a different set of information, perhaps overlapping the first, perhaps not, is relevant to another. If given an index of all items in the catalog, groups would have a hard time deciphering which items were relevant to them. What is needed is a mechanism that defines what items are

relevant to which groups, creates searchable indexes and enables access to those items.

Knowledge maps coupled with metadata models are that mechanism. A knowledge map is a view of the knowledge catalog, allowing the visualization, manipulation, and navigation of the information and knowledge sources contained in the catalog. It uses the metadata layer to derive the connections from resource to resource. Conceptual indexes are abstracted from the complete catalog index, one for each group, enabling organized, disciplined navigation through information. Users are able to customize their knowledge maps, personalizing the information to their needs so that they can make immediate, effective business decisions.

Knowledge is an asset. Like any corporate asset, it can be lost or kindled. Paul Romer of Stanford University identifies knowledge as "the only unlimited resource, the one asset that grows with use." Knowledge can be likened to a living system, growing and changing as it interacts with the environment. Knowledge is created within a corporation through experiences and learning. It is created in the minds of people, as products are created in an enterprise. Ideas breed new ideas, and shared knowledge stays with the giver while it enriches the receiver. It is more valuable than many tangible assets and must be invested in and managed. This concept of successfully managing knowledge many would argue to be the Holy Grail for successful businesses.

Knowledge Portals

Portals can be seen from several perspectives. "Portal" means "large door" or "gateway", indicating that the portal itself is not the final destination but a way to reach many other places. A Web portal is a web site, usually with little content, providing links to many other sites that can either be accessed directly by clicking on a designated part of a browser screen, or can be found by following an organized sequence of related categories. In many cases the required information can also be found through full-text search. Yahoo!, AltaVista, and Excite, are examples of well-known web portals.

Intranet Web Portals

Enterprises are starting to implement intranets, and many are looking at the Web portals trying to apply the lessons learned by the Internet access providers, to satisfy the corporate information access and dissemination requirements. An Intranet Web Portal provides links to all the enterprise relevant sites (internal content providers) and also to some external access providers, such as Yahoo! or HotBot. In many cases, they provide links to external content providers such as Dow Jones or Lexis-Nexis. As with Internet portals, relevant information can be find through extended search facilities or by following a enterprise defined taxonomy, which is usually created by subject matter experts or competency communities, and organized by professional librarians. The extension of the search is limited by usability and technology considerations. These can eventually include not only internal web sites but also databases, file servers, and document management systems.

Because of the increasing load of information, Web Portals provide some personalization facilities that are usually set manually by the users. These basically

define a set of information categories to which the users want easy access, as well as news or changes in web pages they want to be alerted about. The access to content, however, is not unified, as it is in the next point of evolution for enterprise portals: the Information Portals.

Information Portals

Information Portals present users with a gateway providing single logon and single image and a common look and feel, to access the many enterprise sources of content. Users would be provided with all the facilities of an Enterprise Web Portal, but would not have to master, for instance, the intricacies of a query and reporting system in order to access a database or get a report. They would be shielded from the complexity of the underlying tools by the portal.

Information Portals provide a more or less seamless access to information derived from both structured and unstructured data. Business Intelligence tools, such as the IBM VisualWarehouse and Content Management tools such as the IBM Digital Library could be used by a market analyst, for instance, to analyze consumer buying trends while visualizing the characteristics of the enterprise's products and those marketed by the competition. Information Portals could eventually be extended to service the extranet, that is, the net serving the company, its customers, and its suppliers).

The effectiveness of Information Portals can be enhanced by building applications that combine, analyze, and distribute relevant information to the users, and by shifting the focus away from the content of the information to the way the knowledge worker will use it.

Knowledge Portals

Knowledge Portals represent the next step in the evolution of enterprise portals. A Knowledge Portal is the fundamental building block of a Knowledge Management infrastructure. It provides a robust substrate for building a learning organization by providing all the facilities of an Information Catalog plus collaborative facilities, expertise management tools, and a Knowledge Catalog to be used as a repository of institutional memory.

The Knowledge Catalog is a metadata store that supports multiple ways of organizing and gathering content according to the different taxonomies used in the enterprise practice communities, including an enterprise wide taxonomy when defined.

A Knowledge Portal provides two distinct interfaces: a Knowledge Producer interface, supporting the knowledge mapping needs of the knowledge worker in their job of gathering, analyzing, adding value, and sharing information and knowledge among peers, and a Knowledge Consumer interface. This interface facilitates the communication of the produce of the knowledge workers and its dissemination through the enterprise to the right people, at the right time, to improve their decision-making. A sophisticated personalization facility based on user roles as well as habits and preferences are a key feature of Knowledge Portals.

The knowledge workers populate sections of the Knowledge Catalog, and the productivity of this task is vastly improved in a Knowledge Portal through the use of web crawler, concepts, extractor, summarizer, clusterer, and categorizer functions, such as provided today by the IBM Intelligent Miner for Text. These functions can

help discover and make visible relationships in a set of seemingly unrelated pieces of information, providing a context that enables the creation of tacit knowledge or insights.

Collaboration and expertise location are also key pieces of a Knowledge Portal. By integrating the knowledge repository and tools to collaboration environments such as Lotus Notes and Teamroom, the Knowledge Portal will become rapidly the key technology component for establishing an effective enterprise-wide knowledge management discipline.

Critical Success Factors

Implementing a successful knowledge management infrastructure must be approached with a holistic business strategy. Otherwise, enterprises will fall into the same pit that those did in the mid-1980s when they attempted to introduce Total Quality Management in piece-parts, without developing an integrated, strategic plan. The following factors are critical to a successful implementation:

- ?? A well defined project to implement the infrastructure,
- ?? Organizational readiness,
- ?? A well defined taxonomy and ontology,
- ?? A well defined KM architecture with the right technology,
- ?? Organizational culture that embraces knowledge sharing, and
- ?? Strong support from Top Management.

The first component to implementing a successful knowledge management scheme is to define a structured project with a well defined vision, attainable goals and objectives, and quantifiable metrics from which to judge whether those goals were met or not. Examples of metrics are cost reduction, reducing time-to-market for new products, number of new patents and copyrights, increasing sales, and reducing process times. At the heart of any successful venture in the corporate world is a project with dedicated financial and human resources, including executive ownership and a project manager, and this is no exception. While the temptation to put the most junior people onto the task will be strong, understand that the people developing the infrastructure and maintaining it need to know your business model and the organization well.

Organizations must be ready to adopt knowledge management practices before they can be put into use. They must have a clear understanding of the delineation between data, information, and knowledge in order to use knowledge effectively. Many enterprises have mistaken IT tools to be knowledge management tools and then have been disappointed when these tools didn't deliver what they thought they would. Organizations must also realize that while having the technology may assist in the use of knowledge, it does not by itself guarantee knowledge use.

Another critical factor is the ability to define the skill sets within the organization to build a taxonomy or a classification scheme and an ontology for relating and cataloging information. As we said earlier, well-managed information that is properly

cataloged and structured, available and accessible by the right people at the right time becomes knowledge. Different groups within an organization can extract different knowledge out of the same piece of information. A press release announcing the merger of two companies relates different knowledge to the finance group of a competitor than it does to the R&D group. Properly catalogued into a predefined knowledge base, the relevant pieces of information from that single source can be presented to the two groups in a form that would be meaningful to each group.

Fortunately, there are tools out there today to simplify the task of building the classification scheme, also known as the taxonomy. The conceptual components of a KM infrastructure include those tools as well as a repository, also known as a knowledge base, in which metadata is stored, organized and maintained.

This metadata is extracted from internal sources of information such as databases, intranets, document management systems and e-mail repositories and external sources of information, such as news feeds, and other on-line services available on the Internet.

A knowledge management system also includes the components needed to organize, distribute, and visualize information in the proper context to the right people at the right time. Tools here create knowledge maps, pointing the user to the various sources of relevant knowledge. Users are then able to easily navigate through the knowledge and extract the information most meaningful to them.. The map should not only point to documents, they should also point to other individuals in the organization to which the user can go for expertise. Knowledge environment tools extract personnel inventories from on-line Yellow Pages and human resources software already in use.

Knowledge maps are personalized to the user's needs, pulling from the metadata knowledge base only those pieces of knowledge requested from the user. Ease of use is key here--users must be able to query the knowledge base using the language that is natural to them. Also important is visualization--the map must represent pieces of knowledge and their relationships to other pieces in a manner that is quickly comprehensible. For that reason, tools successful in this circle use graphical representations to describe the content and relationships. The old saying, "a picture is worth a thousand words" has never been more applicable. The knowledge environment enables companies to find, organize, and distribute the knowledge needed to make quick informed business decisions.

While technology is extremely useful in the management and distribution of knowledge, it is only successful when the culture in the corporation advocates its use. The culture of today's businesses has not evolve sufficiently to take full advantage of knowledge management. Since knowledge is created in the minds of people through experiences, it can only be captured, preserved, and shared if people are willing to disseminate the knowledge. People rarely give away valuable possessions (including knowledge) without expecting something in return. Even if only partially mindful of doing so, people make choices about how to spend their limited time and energy and base those choices on perceived self-interest. In an era when knowledge is power and premiums are paid for that knowledge, free dissemination of knowledge is unlikely.

In the 1997 Journal of Knowledge Management survey, organizational culture was seen as the biggest obstacle to creating a knowledge-based organization. Eighty percent of the respondents reported that their organization's culture either actively or passively hindered the development and introduction of knowledge management strategies and programs. This finding was valid for all industries/sectors and all geographic regions. In order for businesses to capitalize on the power of shared knowledge, they will need to alter their corporate culture to reward those that disseminate their knowledge freely to the corporation at large. One example of how that might be accomplished is to make employee contributions to the company's knowledge base a formal entry in the standard employee performance appraisal system. Another example might be through the use of open, electronic bulletin boards for problem solving--those that contribute to the bulletin boards are praised--those that don't, especially managers, become very conspicuous. Everything is done in the open, with no place to hide, and that leads to people sharing as everyone in the corporation can see the value of an individual's contribution.

Companies that do understand the power of shared knowledge sometimes will not take the step to implement a knowledge management system on the basis that sharing knowledge is everybody's job. They should not need to "waste" resources on developing, training, maintaining, and promoting the use of such a system. However, knowledge management will not succeed without workers and managers whose primary jobs involve extracting and editing knowledge from those who have it, facilitating knowledge networks, and setting up and managing knowledge technology infrastructures. Top management needs to understand this and be willing to make significant organizational changes to support the effort, putting their best employees on the project. Complete organizational mind share is required and initiatives from upper management are the most effective method of obtaining that mind share. Managerial influences are based on leadership, coordination, control and measurement. Of these four elements, leadership is the primary factor in determining how successful the knowledge management implementation will be.

Conclusion

In this new world of information highways, the roadmap to successful knowledge management has never been more needed. Data abounds and the skill is knowing how to make use of it. Technology innovations have propelled us into the information age, bringing more information to our desktops than we have time to sift through. Companies are going global, with groups of people scattered all over the planet. Competition is fierce, causing shortened time-to-market windows and requiring more flexibility in product offerings. And of course, shareholders are demanding more profitability.

Information technology has kept up with the pace in terms of storing the data and making it readily available to the organization at large, however, to truly capitalize on the expertise latent in an organization, a knowledge management infrastructure needs to be employed to harness all the technologies together, process the information, organize the knowledge, create a knowledge map personalized to each user, and make the knowledge sharing process easy enough that they keep coming back to the well.

Tools in and of themselves will not transform a company into a knowledge machine. The culture of a company must encourage and reward knowledge sharing, and get away from the "knowledge is power" syndrome that inhibits knowledge transfer. Knowledge sharing must be built into workloads, giving employees time to share their expertise. Above all, the abilities and experiences of every employee must be valued in order for trust to be kindled.

Knowledge management requires a shift--a shift in the way we look at information, a shift in the IT infrastructure we use, a shift in corporate culture, and a shift in our workflow. It is an investment and like all investments, requires diligence and time to return it's potential value. Companies that have endorsed the knowledge management concept and have put an infrastructure to work are realizing the benefits--they are:

- ?? Using actionable knowledge to gain and maintain competitive advantage,
- ?? Making more informed business decisions, and making them more quickly,
- ?? Uncovering hidden opportunities and capitalizing on them more easily,
- ?? Spending less time searching for the right information,
- ?? Spending less time recreating the work already done by others, and
- ?? Improving internal communications through easy access to enterprise knowledge.

Knowledge is a valuable corporate asset that must be fostered, invested in and managed. With the right tools and the right attitude, companies can leap from the information age to the knowledge age and come out on the other side more profitable and better prepared for the future.

CASE STUDY 1: Competitive Intelligence

The senior managers of your company are establishing the long-term business strategies of your company, and deciding on some tactical investments to support these strategies. At the next board meetings, the following issues will be analyzed: How does your company compare to its competitors? Which companies are most active in your company's area of business? What are the most promising new technologies related to your company? What is going on in related areas of research? Which governments are subsidizing the development of this technology?

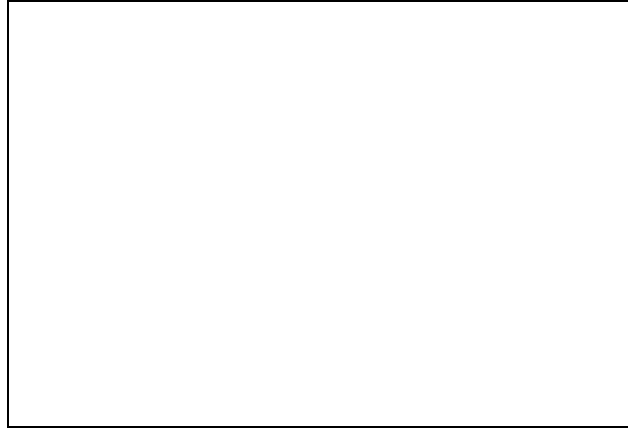
How does this information relate to your company's international operations? As the company CIO, you need to provide actionable knowledge for executive decisions. How can a knowledge management infrastructure help?

A large part of the competitive information is publicly available in the form of text. Related patents, reports, newsfeeds and web pages provide an enormous volume of potentially useful material that is very hard to digest. For instance, one global company in the electronics industry that planned to establish manufacturing operations in Korea needed to understand the trends discovered in the research done by the local companies to determine which areas were likely to be subsidized by the Korean government. A preliminary survey indicated that in the last year, 3881 patents were granted to the companies in the electronic industries. The printout of the patent reports was stacked in a pile of paper four feet high. To read these reports and obtain meaningful conclusions in the required time frame seemed to be out of the question. The company used the text analysis technology available in the IBM Intelligent Miner for Text to solve this problem. The approach taken was to discover main areas of research by identifying clusters of patents -- patents that have common themes. In a short time the team was able to identify thirty-three main research subjects, ranging from semiconductor manufacturing to LCD devices and multimedia enhancing hardware, as shown in Figure 1.

| 3881 PATENTS - 33 Subjects | |
|-----------------------------------|--------------------------------------|
| Subject 1 | Semiconductor manufacturing (201) |
| Subject 2 | Memory devices (154) |
| Subject 3 | Messaging (119) |
| Subject 4 | VCR and tape recording hardware (98) |
| Subject 5 | TV output control (71) |
| Subject 6 | Image transmission (63) |
| Subject 7 | Optical disc technology (57) |
| | |
| Subject 31 | Computers (25) |
| Subject 32 | LCDs (25) |
| Subject 33 | Multimedia (25) |

Figure 1: Text Mining of Patent Reports

The next step was to present a meaningful map of the companies involved in each research area, indicating the number of patents awarded to each of them. Using the Intelligent Miner visualization functions, the clusters appeared as shown in Figure 2.



The black lines indicate contextual links that represent relationships discovered during the clustering process. The graph displayed on a computer screen using this tool is an active image: by clicking on each cluster the user can have access to the list of patents that comprise each cluster, and by clicking on a patent on that list, the text of the patent is displayed.

Knowledge Management for Competitive Intelligence

Knowledge discovery through text analysis is an important knowledge management technique. It helps the corporate analyst by finding important relationships that can foster new insights through clustering techniques, or simply by categorizing the information sources according to a useful corporate taxonomy. The analyst can thus find together all the sources relevant to the aspects of the competition or the market place she is interested in. Using the IBM Intelligent Miner for Text, the categorization of the documents is done automatically by the system after it is trained with samples from the desired categories. Although clustering can be helpful to discover possible categories, large enterprises find it more productive to use their corporate-developed taxonomy as the categorization schema.

Knowledge mapping is another key competitive analysis technique. A knowledge map used for competitive intelligence provides a clear and easy-to-understand view of a competitor's organization. The knowledge map shown in Figure 3, built with IBM KnowledgeX, shows a series of nodes and links.



Figure 3: A Competitor's Knowledge Map

The nodes are the key business elements of the competitor organization, gathered from the press releases and analysis documents available on the Internet. The links represent relationships among the objects representing the business elements, with each object playing a defined role in the relationship. Each object can have an associated behavior: when clicking on the icon representing a document, we may open the document using a word processor, or link to the web address (URL) where the document is stored. When clicking on a production, we can get a photography of the product, a catalog page or a set of technical specifications.

People are important business elements and can be represented on a knowledge map. We can therefore build maps describing the boards of directors of several competitors, associated businesses and holdings. IBM KnowledgeX allows us to discover unexpected relationships such as "who are the directors of our sister company that are also seated on the boards of subsidiaries of our competitors," through simple drag and drop of knowledge maps.

Competitive Intelligence is therefore another aspect of business knowledge, where presenting the right information in the proper context, distilled and related to other sources of information and knowledge, helps a manager to get the insights to define the adequate market strategy or to act with the appropriate tactics. A solid knowledge management infrastructure supporting the organizational learning process allows a company to leverage what it has learned and be a winner in the marketplace.