

A first proposition for a multi-dimensional classification-based system for corporate knowledge management

Laszlo Szathmary and Amedeo NAPOLI
LORIA -- UMR 7503
BP 239, 54506 Vandoeuvre-lès-Nancy Cedex, France
Email : {Laszlo.Szathmary,Amedeo.Napoli}@loria.fr

Abstract

This paper presents an ongoing research work on knowledge management. At present, information systems are complex, consisting of many different elements, and are increasingly overwhelming users in the range of information they support. One area in which enterprise portals can provide considerable value is knowledge management. The paper is organized as follows. In the first part, we present the state of the art of the needs for knowledge management in an enterprise. Then, we present our proposition for organizing knowledge within a multidimensional referential. We propose an example and then we conclude the paper.

Keywords : viewpoints, object-based representation, multi-dimensional referential, knowledge-based system

1 Introduction

The primary goal of knowledge management is to improve organizational performance by enabling individuals to capture, share and apply their collective knowledge to make optimal decisions. About 80% of the world's biggest companies have knowledge management efforts under way [RGS00].

In the knowledge-powered enterprise – which is yet an "ideal enterprise", not a real one – the knowledge management happens in the background. It is done by everyone as part of the day-to-day job, embedded in the workflow. People are easily able to obtain the data, information and knowledge they need to do their jobs. They interact effectively with their colleagues – anywhere, anytime and by any means. The company is able to connect the people with the knowledge to the people who need it. In the knowledge-powered enterprise, knowledge sharing and application are standard. As a result, the enterprise can improve its performance and reduce its costs [RGS00].

The problem is to be able to design a "multidimensional portfolio," in order to organize and to index the different aspects of knowledge in an enterprise. We have chosen an object-oriented view

for such a representation. Every piece of knowledge is represented within classes in an object-based representation system. The multidimensional organization is supported by a number of tangled hierarchies of classes, one hierarchy giving a particular point of view on the enterprise knowledge.

2 State of the art

Why is knowledge management so picked up today ? It has several reasons :

- (1) the collaborative space : it is now more virtual than physical. Companies are increasingly distributed worldwide.
- (2) the intellectual capital : in this age of information, knowledge is the most important factor.
- (3) Information technology (IT) : as a result of IT, it is now possible to do something about knowledge management. Today's internet, intranet and web technology permit practical capture, sharing and leveraging of information and knowledge throughout organizations [RGS00].

The tasks of knowledge management are neverending. Knowledge is never fully managed. One reason of this is that the categories of required knowledge are always changing. New technologies, management approaches, regulatory issues, and customer concerns are always emerging. Companies change their strategies, organizational structures, etc. New managers and professionals have new needs for knowledge [THD98].

Technology

An organization must have "good enough" technology to make progress. It isn't everything, but without this little progress will be made [RGS00].

Knowledge management is expensive, but not managing knowledge is even more so. How much does it cost an organization to forget what key employees know, to not be able to answer customer questions quickly or at all, or to make poor decisions based on faulty knowledge [THD98]?

It is important to get started as soon as possible. Begin to build the community with the available technology and upgrade it later.

- (1) Determine what technology the community members actually use. For example they use e-mail every day, and web almost every day.
- (2) Start the communities off with a small evolution to the technology they already used [RGS00].

Building an intranet portal is today a standard first step in knowledge management.

One central problem is to provide the right information at the right person at the right moment. This is one question that has to be answered in a corporate knowledge portal. Moreover, other questions are related to the preceding : how an employee may have access to the needed information ? How this employee can know who is the right contact to get the information ? What are the means that can be exploited to insure a correct knowledge sharing ?

A portal is a single point of access to knowledge ; a kind of "reference point" through which information may be available. A portal is an application that gives users a single gateway to the information and applications they need to do their jobs. It draws together on the desktop all the important information from both inside and outside a company [RGS00]. A portal provides unified access to all the organization's information, both unstructured and structured.

Unstructured : mostly web pages and documents ; structured : usually stored in databases.

Portals deliver information to the desktop by way of a "thin client", which is usually a web browser [RGS00].

But building the infrastructure is only the beginning. Access is important, but successful knowledge management also requires attention and engagement. In order for knowledge consumers to pay attention to knowledge, they must become more than passive recipients [THD98].

A portal gives access to:

- bulletin boards : they are like newsgroups. The e-mail messages are captured in web-searchable archives
- documents : training materials, product information, technical manuals, policies, etc.
- workflow : this is the series of steps to be followed to perform particular tasks
- software : for example job planning, simulation, diagnosis
- project archive : (1) the archive improves the efficiency of project teams. It also helps to bring new team members up to speed more quickly by giving them access to the historical record. (2) It reduces the "knowledge drain" as people move around or leave the company.
- expertise directory : for example the LDAP (Lightweight Directory Access Protocol , "information for use")
- best practices, lessons learned
- news
- FAQs [RGS00]

Portals typically offer two ways to find information :

1. search : it can be full text or context specific ; indexing of structured and unstructured data from file systems, web servers, e-mails, etc.
2. browsing through a hierarchy of categories : creation and maintenance of a common vocabulary, and one or more category hierarchies are key tasks associated with portal construction. A recent survey shows that 10% of people find information by browsing, 10% by searching and 37% by both searching and browsing ; so it is crucial to design portals that support distinct information-searching paradigms. There is need for content-based document search. Using the metadata associated to the documents, it is possible to create programmatic links to specific documents, all documents of a particular document type, or all documents with a particular attribute and value. This creates a more "hands-off" publishing scheme for the Intranet, as the technical staff doesn't need to be involved on a day-to-day basis, and on the other hand it is easier for the users to browse among the documents if several categorizations are supported [PGD01].

3 Our proposition

We present our proposition for a "multidimensional portfolio for corporate knowledge organization and indexing." An example of a primitive architecture for a prototype knowledge management problem is shown in Figure 1.

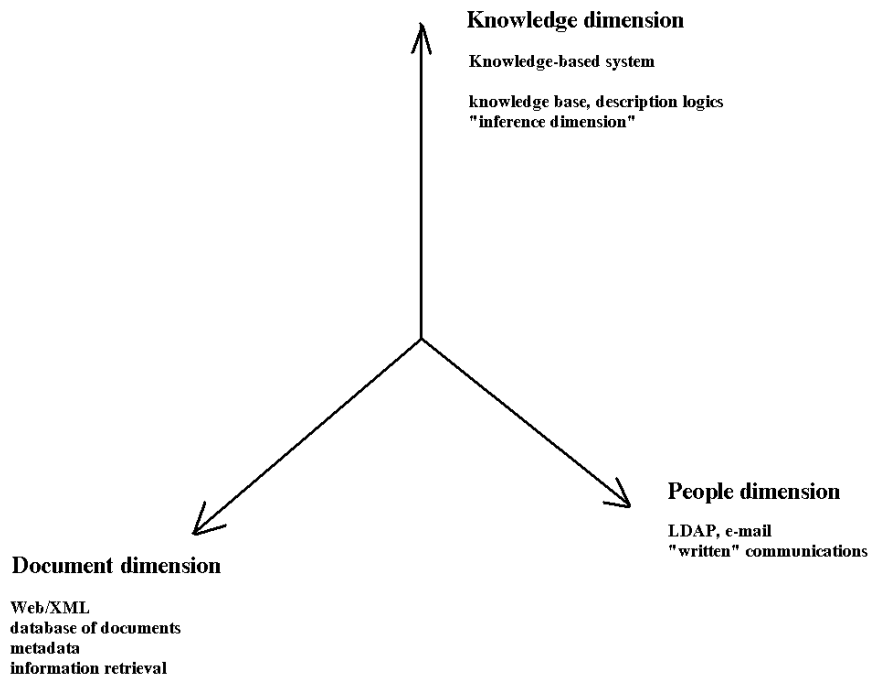


Figure 1. : primitive architecture

In the architecture we distinguish three dimensions : people, documents, knowledge. Let's see them in details :

Dimension of people

It is the people who work at a company and whose work have to be helped by the knowledge management. One basic organizational unit of knowledge management is the Community of Practice (CoP). It is a small group of people who have worked together over a period of time. There are many CoPs within a single company, and most people belong to more than one of them. A primary focus of knowledge management work is finding effective ways to connect groups of people [RGS00]. Today, when the collaboration space is more virtual than physical, significant attention must be paid to the people. When the members are spread around the world in remote areas, their work has to be coordinated. "Does the left hand know what the right hand is doing ? "

The greatest difficulty is to change people's behaviour. Organizations must find ways to motivate individual community members to share what they know and to apply the knowledge of others [RGS00]. Knowledge management without participation will fail. (There are some firms that are beginning to evaluate and reward personnel for knowledge sharing and use) [THD98].

Here are two examples of CoP operation :

- (1) e-mail : this is the tool of choice to support CoP interaction
- (2) Web : since 1993 the web has become the primary means for finding information.

These tools are used by everyone and thus they have to be supported [RGS00].

Knowledge champion : he is responsible for animating the community, encouraging the members to participate, highlighting successes, etc. [RGS00] He edits the community news. He is responsible for the success of the portal. If the firm is bigger then not only one person but some group will be required. Among the tasks that such a group might perform are :

- collecting and categorizing knowledge
- establishing a knowledge-oriented technology infrastructure
- monitoring the use of knowledge [THD98]

The community members should be able to find each other by searching the corporate directory. "Know who" is often as important as "know what" and "know how". Example : the LDAP. The directory contains the basic "coordinates" of each employee : telephone number, e-mail address,

physical address, business unit(s), the manager(s) to whom they report, photographs, pointers to their personal pages, etc. The employees are responsible for keeping their information up to date. However this is not an easy task, because people tend to forget to update their data. It is certainly a hard problem to automate this kind of task, and one way to manage such a problem is the following: when people last changed their information should be logged, and if it exceeds a limit (for example one month) they are sent a reminder message. The update process has to be kept as simple as possible – a web page where users can modify the fields is perfect. It is very important to keep the LDAP up-to-date, because this way community members can find others who might be able to offer assistance. If it contains old information it will become unusable.

Dimension of documents

Documents are stored in databases with the appropriate metadata. Today, and in the future, XML is emerging as the format of choice for a variety of types of data, especially documents [EXD00]. XML makes searching simpler and more dynamic. XML enables the extensive reuse of material. This allows enterprises to turn the same content into press releases, white papers, brochures, presentations, and Web pages through the use of XSL stylesheets. In addition, XML can handle all kinds of data, including text, images and sound, and is user-extensible to handle anything special. The problem until now has been how to store the XML-tagged data. A promising solution is to use databases [EXD00]. There are two possibilities. The first is when we use existing database types, and existing database products. Its advantage is that ordinary relational, and OO databases, are well known, users are familiar with them, and they are a safe choice in the corporate mind. Its disadvantage is that translating XML to and from the database requires considerable processing, especially for large and complex documents. The second solution is when we store XML in its native form, thus the translation between XML and the database is not necessary. Unfortunately these databases are not well known yet [EXD00] ; the design of such DBMS is in progress. Another rather quick and dirty solution is to index XML documents and to store indices in a database. The problem is that the content-based access to documents may be lost.

It is the responsibility of the knowledge champion to maintain, organize the documents. When the intranet portal opens for the community, there must be a possibility that the members find something of value and hence be encouraged to return to the site in future and to contribute to its evolution. Document management is the way in which a business creates, revises, publishes, and purges electronic information assets [PGD01]. Currently, in most companies, documents are stored in a number of different places : individual PC hard drives, public and private network drives,

disks, web servers, zip drives, paper documents in file cabinets, etc. This results in documents being stored in multiple places and islands of unorganized information that is difficult to manage. The result is being unable to find information, potentially using out of date or inaccurate information, and limiting the ability to leverage existing information. With a document management system these problems could be solved. A document management system can coordinate the changes, access, and availability of business critical information on a global scale. The most current version of any document is always the one that is presented so users know they are using the "right" document. In addition a version history is available to provide access to previous versions from any point in time during the document life cycle. Searching for information can be simplified by using metadata [PGD01].

What kind of documents are there in a portal ?

1. "Normal" documents, written by an employee. The knowledge has to be captured at its source ! Metadata also have to be added to make the retrieval easier.

(1) implicit metadata (contributor, date, content-type)

(2) explicit metadata (titles and descriptions, user-definened types and attributes) [TG00]. If a user modifies a document, he has to write in a few lines what and why he has edited. This information is added to the metadata. After modifying a document, the new document will be the up-to-date (newest) version, but the older ones are still available in the archive – we need versioning. If somebody searches this document, he will find the latest version. The system should give the ability to subscribe to newsletters that inform when a new document was added (or edited). After downloading a document a user can rate that. When searching documents this information is also shown, and thus users can choose according to the rates too.

It is not a good idea to dump all the documents and sort them later. Knowledge must be captured in context, so metadata can be used here to sort the documents automatically.

2. Best practices. How can a knowledge champion put initial content to the portal ? For example, he examines the past year's messages. He finds messages that summarizes best practices and seeds the knowledge hub with them [RGS00]. Later users will send best practices. When a question was posed to the bulletin board, a number of responses received, and the question eventually resolved to the satisfaction of the members. One of them would write a summary message, detailing what had been decided. Placing these summaries on the knowledge hub in the appropriate categories improves the ability of the members to find them quickly and therefore increases the likelihood that they would use them [RGS00]. As a research work on

that subject : a kind of text generation process here is needed, able to propose automatically a survey or an abstract of a number of texts relative to a given point.

3. Lessons learned, experiences. Contribute lessons learned after the job has been completed. It is the most difficult thing : to create a sharing culture. This is the same problem as creating an ontology or a knowledge base in a given domain.

If the portal is started containing useful information for the members and if they can do successful search the impetus for adding knowledge will increase. With everyone searching the internal knowledge base, a critical mass of knowledge and perception will develop to create a virtuous cycle. The more knowledge in the knowledge base, the higher the chance for a successful search [DMa].

How can employees add documents to the knowledge hub ?

Busy people with "regular day jobs" are the front-line knowledge management practitioners. These people build the knowledge hub and they are not computer specialists, so the tools must be extremely quick and easy to use ; they must be embedded in a user's normal workflow [RGS00].

- A simple solution : replace the standard Save command with an applet that automatically prompts the user to see if he would like to add the document to the knowledge base. If the employee agrees, the document is saved to both the file location of his choice, and the knowledge base with the click of a button [DMa]. The user must be able to add metadata to the document, but this is not an easy task.

Search

Search technology has become a crucial part of today's connected world. In knowledge management or intranet applications, the primary focus is on accessing internal resources (corporate file systems, databases, document repositories) and external resources for internal purposes (for example newsfeeds or Web sites) [APSa]. The intranet portal has to support both full text and metadata search on the internal documents.

Unfortunately most users cannot accomplish effective search. They usually use single search term ; multiple search terms are seldom and Boolean compositions are rare. At most companies there are no satisfactory education programs teaching the employees how to use the available tools. Information agents can help here, who interact with the search engine and create more requests using synonyms of the given term [CL00].

Dimension of knowledge

Knowledge is broader, deeper, and richer than data or information [GS00]. Knowledge for an employee means organized, applied information, that is a justified true belief [TR01]. In the knowledge-based system knowledge is syntax, semantics, and inference.

Knowledge management needs AI, for example :

- (1) Knowledge Base : different words can be used, corporate memory, knowledge repository, best-practices database, etc. A high-quality knowledge base is fundamental for successful knowledge management. Ontologies are needed. It is also very important to have reasoning formalisms for decision helping, and in particular, case-based reasoning is one of these formalisms. We have decided to investigate this research line in the next step.
- (2) Search : community members must be able to find and bring to bear the relevant knowledge in the repository. Agents and distributed problem-solving technology will play an increasingly important role. Moreover, but this is not already the subject here, data mining tools are also needed.

4 An example

Let's imagine a software company with customer service. The contents of the internet and intranet portal are made from the same data sources (databases, knowledge bases), the only difference is that the intranet portal can show more information (privileges). This way a customer and an assistant can look at the same page if required. In the company there are several CoPs. Each project has a homepage, containing the project description, the list of the developers, the project history, etc. Each project has a mailing-list, whose archive is available on the web. If an employee needs help, he can search in the LDAP directory. When a user writes a document and saves it, the system asks if he would like to send it to the document base too. If he chooses yes, he has to give some metainformation, which makes the further search easier. Behind every project page there is a person – who is a project member too – who is responsible for managing the homepage (for example next meeting, and so on), and behind the whole portal there is a group of people, who are responsible for collecting, organizing any data, document that is sent to the knowledge base. They are animating the intranet, they are responsible for its success. Documents are stored in databases, the majority is in XML.

5 Conclusion

Enabling organizations to capture, share and apply the collective experience and know-how of their people is seen as fundamental to competing in the knowledge economy. In this paper we have given a possible multidimensional classification showing the connections between people, documents and knowledge. This is a preliminary work to present the principles on which we plan to design a corporate knowledge management system. Similar work is done at the ACACIA research team at INRIA-Sophia-Antipolis, aiming at offering methodological and software support in form of models, methods and tools for knowledge management, ie. for building, managing, and distributing a corporate memory [RD00].

In this age of information, knowledge is the most important factor in the long-term success of both an individual and an organization.

Bibliography

- [APSa] : Alan Pelz-Sharpe, The next generation of search,
http://www.kmworld.com/publications/magazine/index.cfm?action=readarticle&article_id=1154&publication_id=1
- [CL00] : Christopher Lueg–Reinhard Riedl, How Information Technology Could Benefit from Modern Approaches to Knowledge Management, PAKM2000, http://SunSITE.Informatik.RWTH-Aachen.DE/Publications/CEUR-WS/Vol-34/lueg_riedl.pdf
- [DMa] : Daryl Morey, Knowledge Management Architecture,
http://www.brint.com/members/online/120205/kmarch/{kmarch_1.htm,...,kmarch_7.htm}
- [EXD00] : Edmund X. DeJesus, XML Enters The DBMS Arena, Oct. 30, 2000,
http://www.computerworld.com/cwi/Printer_Friendly_Version/0,1212,NAV63-129-1375-1382_STO53026-,00.html
- [GS00] : Guus Schreiber, Hans Akkermans, Anjo Anjewierden, Robert de Hoog, Nigel Shadbolt, Walter Van der Velde, and Bob Wielinga ; Knowledge Engineering and Management : The CommonKADS Methodology, 2000
- [PGD01] : P. G. Daly, Why Document Management? , Intranet Journal, 2001,
http://intranetjournal.com/articles/200108/pit_08_08_01a.html
- [RD00] : Rose Dieng, Corporate Knowledge Management through Intranet and Internet, 2000,
http://www.ercim.org/publication/Ercim_News/enw41/dieng.html
- [RGS00] : Reid G. Smith and Adam Farquhar, The Road Ahead for Knowledge Management : an AI Perspective , AI Magazine. Volume 21, 2000, Number 4, p.17–40

[TG00] : Thomas Gruber, Collaborative Knowledge Work : Theory and Practice of a Successful Commercial Application, 2000, <http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-34/gruber.pdf>

[THD98] : Thomas H. Davenport, Some Principles of Knowledge Management, 1998, <http://www.bus.utexas.edu/kman/kmprin.html>

[TR01] : Tom Reamy, From Information Architecture to Knowledge Architecture, September/October 2001, http://www.infoday.com/IP/sep01/reamy_intro.htm