Knowledge Management and the Internet

Rose Dieng, Inria

WHAT DO WE KNOW AND HOW can we manage our knowledge assets to fulfill our organizational objectives? These are important questions for an organization. Knowledge management should enhance individual, group, and organizational learning; improve information circulation; and even support innovation. It aims to capture and represent an organization's knowledge assets to facilitate knowledge access, sharing, and reuse. In doing so, it attempts to address these possible goals: capitalizing on individual know-how in a collective knowledge; improving newcomer learning and integration; disseminating best practices; improving corporate work processes, product quality, and productivity; and reducing new product design times. We must tackle these complex problems from at least the human, socio-organizational, and technical viewpoints.

A corporate or organizational memory is an

> explicit, disembodied, persistent representation of crucial knowledge and information in an organization, in order to facilitate their access, sharing, and reuse by members of the organization, for their individual or collective tasks.[1,2]

That organization can be an actual enterprise, a public organization, or just a department or service. It can also be a group, community, or virtual enterprise comprising members from different companies, gathered by a common interest.

Different scopes and grains are possible for a corporate memory, but its development relies on these steps:[2] detection of needs, construction, diffusion, use, evaluation, and maintenance and evolution. We can model the organizational memory from several perspectives; for example, John Kingston and Ann Macintosh emphasize the what, how, when, who, where, and why perspectives.[3] Fundamentally, an organizational memory aims to deliver the right knowledge to the right person at the right time in the right format to enable the right action.

Techniques for building an organizational memory can be noncomputational, database-, document-, knowledge-, case-, or Web-based;[2] they can also be product- (based on repositories) or process-oriented (based on corporate work processes). The choice depends on the type of organization, its needs, and its culture and must take into account people, organizations, and technology. (See the sidebar for a description of some of these techniques and articles in this special issue that use them.)

Exploiting the Internet and intranets

Individuals and organizations can take advantage of the remarkable possibilities of access to information and knowledge that the Internet provides. Web technologies such as HTTP and HTML have dramatically changed enterprise information management. Moreover, an intranet relying on Internet technology and protocols enables intraorganizational communication and internal information sharing through the corporate internal network. For example, a multinational corporation can benefit from intranets and the Internet to gather, manage, distribute, and share knowledge, inside and outside the corporation. Figure 1 shows the role of the Internet and intranets in the corporate memory management cycle.

Corporate memory creation and revision

A company can exploit the Internet and intranet features in several ways. It can use internal HTML or XML pages or external URLs containing organizational memory, making it accessible throughout the company. Internal or external discussion forums, newsgroups, news archives, and FAQs also encourage information exchange.

More proactive methods of creating and revising corporate memory include integrating messages exchanged through e-mail in the corporate memory, extracting information from the external Web sources for technological or strategic intelligence, and using computer-supported cooperative work (CSCW) tools to support complex-system collaborative design or collaborative software development.

The wide variety of organizational choices involves several actors with different roles:

- human knowledge sources (such as experts, specialists, or operators), whose knowledge must be made explicit or who

1094-7167/00/$10.00 © 2000 IEEE

have written documents that others will access through the organizational memory;
- knowledge engineers, who acquire and model knowledge;
- knowledge watchers, who gather, filter, analyze, and distribute knowledge elements from the external world (from external information Web sources, for example);
- organizational memory developers, who concretely build, organize, annotate, maintain, and evolve the corporate memory;
- a team of validating experts (for example, a reference team), who validate the knowledge elements before their insertion in the organizational memory;
- corporate memory users, who must easily access and reuse memory elements; and
- organizational memory managers, who supervise the organizational memory project.

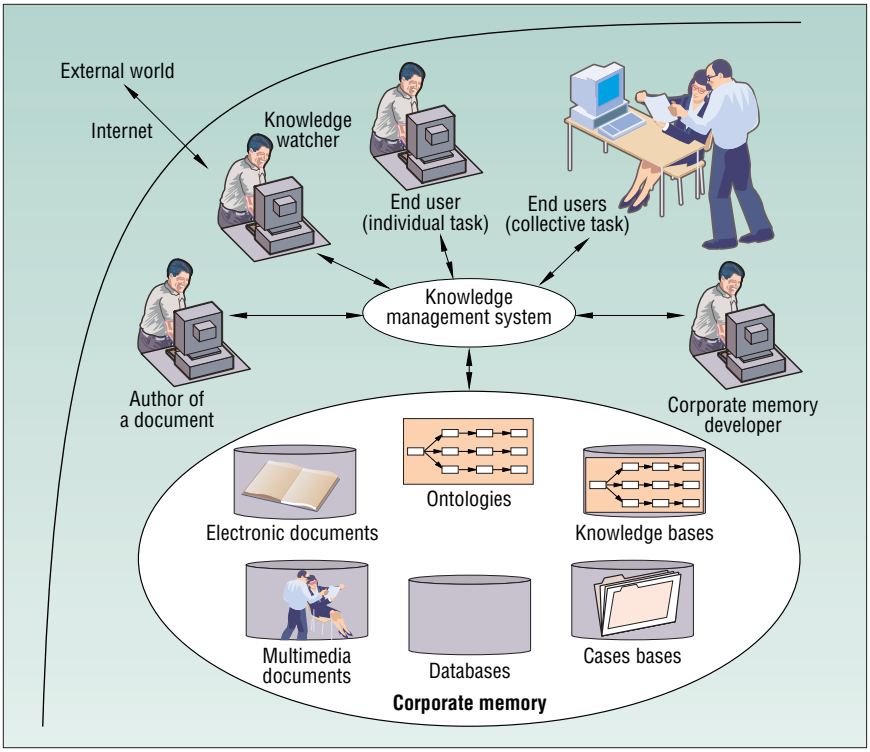

Figure 1. Exploitation of the Internet and intranets in the corporate memory management cycle.[2]

External information sources. A corporate memory can rely on the internal competencies inside the company and on external information sources that provide knowledge from the external world—that is, useful for enterprise activities. Knowledge watchers collect scientific, technical, and economic information from journals, newspapers, and the Web. They then filter, analyze, and validate that information to integrate some interesting elements into the corporate memory or forward them to the relevant members of the company. An intranet can exploit an internal corporate memory, while an external memory can rely on either an extranet connecting the company and some privileged partners such as customers, suppliers, and subcontractors or the Internet and the Web.

The different kinds of intelligence interesting to a company include

- *technological intelligence*, to follow an existing or an emerging technology;
- *competitive intelligence*, to know about activities, products, or services of competitors or other actors in the enterprise market;
- *commercial intelligence*, to know about the enterprise commercial environment such as distributors, suppliers, and customers; and
- *strategic intelligence*, to support the enterprise managers' strategic decisions.

Agents monitoring news or Web information can gather and filter information for technology watching. Companies can also exploit research on intelligent integration of information from possibly heterogeneous, voluminous sources of information on the Web.

Corporate memory organization. In the framework of an organizational memory materialized in documents (or informal knowledge), we can associate such documents with a formal knowledge on which we can perform a reasoning to retrieve adequate parts of the document. This formal knowledge can either represent a part of the document or consist of metaknowledge about the document—possibly with more knowledge than the document includes. (Several articles in this special issue use this approach; see the sidebar for examples.)

For example, Ontobroker exploits an ontology to guide information retrieval from annotated HTML documents accessible on the Web.[4] This approach aims to improve classic Web search engines with semantics-based information search capabilities and relies on semantic metadata or annotations on the resources constituting the corporate memory. Some research on text-mining techniques aims to automate partly the building of such semantic metadata, using textual-document automatic analysis based either on statistical or linguistic techniques. Recent research on the semantic Web[5] can have a significant impact on corporate knowledge management.

Dealing with documents in multiple formats is often needed because not all companies impose a uniform document format. In this instance, storing short document descriptions (or semantic metadata about them) with hypertext links in the organizational memory instead of the whole documents is a possible solution.[6] Another solution is to wrap company documents in XML documents. Moreover, ongoing research on the Resource Description Framework could let us express semantic metadata on the resources constituting the corporate memory.

In the future, thanks to XLink, we will be able to exploit sophisticated hypertext links. We can implement different types of presentation of the same document, which lets a company present an XML-based corporate memory differently according to user profiles. For example, the disciplines, professions, or roles of the project participants or the interests and contexts of the potential users can determine organization memory views. However, organizing the corporate memory through static views related to the knowledge sources and potential user profiles might be insufficient. A knowledge management system can also dynamically create the corporate memory view when a user tries to access the organizational memory in a given context, provided that the system can access a user's profile or a task description.

Special issue articles

This special issue presents several approaches, including e-mail, contexts and views; ontology-guided knowledge retrieval; and collaborative design over the Internet.

E-mail, contexts and views

In "Tying Knowledge to Action with kMail," David Schwartz and Dov Te'eni exploit the Internet and e-mail to disseminate knowledge. They emphasize the need to tie knowledge to action and present the kMail system, which integrates e-mail with organizational memories to distribute knowledge over the Internet at the right time and to the right person. kMail lets e-mail senders contextualize their messages by creating relevant views of the organizational memory. It also uses metaknowledge consisting of user-profile information and shared semantics information. They describe the context through such metaknowledge and link it to the organizational memory each time knowledge is used to perform an action. Domain-specific criteria help determine the relevant views in the organization.

Ontology-guided knowledge retrieval

John Domingue and Enrico Motta's article, "PlanetOnto: From News Publishing to Integrated Knowledge Management Support," presents a set of tools, PlanetOnto, enabling the members of an academic community to build collaboratively and share a common archive of news items. PlanetOnto lets the user associate an archived document to an ontology and allows semantics-based information search and knowledge retrieval. It includes agents providing personalized news alerts and identifies news potentially interesting to a given user. The tool also lets users enrich a document by associating it with a formal representation, supporting various kinds of reasoning.

Philippe Martin and Peter Eklund present WebKB, a Web-accessible tool that interprets semantic statements stored in Web-accessible documents, in "Knowledge Retrieval and the World Wide Web." They express these statements in a precise, flexible, and scalable formalism: conceptual graphs. WebKB lets the user index any element of a Web-accessible document with such knowledge statements. The relationships connect elements in one or more documents. WebKB offers query commands and a script language,

Cooperatively creating and revising. Several members of a company can use the Internet to cooperatively create and reuse organizational memory. In this case, distributed editions of the corporate memory must be possible and the consistency and security problems arising from this distant edition must be solved.

According to the organizational choices, the corporate memory creation and evolution can be distributed (for example, each knowledge source can directly update the memory) or centralized (for example, by a person or a service that manages the corporate memory). A distributed corporate memory can support collaboration and knowledge sharing between multiple people in an organization or in several collaborative organizations, even if they are geographically dispersed. It is particularly useful for dynamically building and preserving the memory of an ongoing project. The explicit representation and visualization of argumentations exchanged during team discussions let the team members collaboratively capture an ongoing project memory.

A particular case of project memory is a design project memory. Companies can use these techniques to keep a memory of software development projects and of complex system design projects and might let designers exchange knowledge both on the design products and on the design processes.

Ontologies. Because a corporate ontology might be part of an organizational memory,[7] companies can exploit tools for cooperatively building a consensual ontology.[8] Examples of such tools are the Ontolingua server,[9] Apecks,[10] and WebOnto.[11] Ontologies can help index the organizational memory and to enable later semantic searches and knowledge retrieval in the corporate memory materialized in documents or in news archives. They are also useful for pushing relevant information to the relevant users. When the organizational memory is materialized in a knowledge base, tools supporting knowledge acquisition, modeling, and inference through the Web, such as WebGrid, become important.[12]

Knowledge diffusion

Knowledge diffusion lets enterprises take advantage of popular Internet services such as e-mail and the Web. The Web can be a basis for uniform information distribution, independent of how you store information. Diffusion can thus rely on populating knowledge elements on the Web or on a knowledge server on the Web.[8,13,14] Users can access documents (such as HTML and XML documents), relational or object-oriented databases, ontologies, knowledge bases, case bases, digital journal articles, and news archives over the Internet. Therefore, companies should consider multiple knowledge servers: document, ontology, knowledge-base, database, electronic journal or digital library, or news.

The main problems are

- organizing and possibly indexing the corporate memory to enhance its diffusion;
- retrieving relevant elements of the corporate memory to answer a user's request or proactively push relevant elements toward users; and
- adapting the answer to users, in particular to their tasks, according to the corporate work processes.

Such problems seem similar to those of information retrieval from the Web, but they are specific to an organizational-memory framework. Therefore, solutions such as Web search engines or Web intelligent agents might be useful but must be adapted in the context of a corporate memory. For example, they can rely on a domain ontology or an enterprise model to improve information search.

In addition to the numerous innovative applications of the Web,[15] we must take into account human factors because using the Web involves new types of human–computer interaction and human–human cooperation.[16] In some companies, an intranet can be the tool for a corporate reorganization, aiming at a less vertical and more horizontal organization, because it can enable information distribution at several (or even all) hierarchical levels.

One goal of knowledge distribution is to improve individual, group, and organizational learning. Several kinds of members of a company might be interested in knowledge distribution, such as managers, designers, engineers, and employees.

Knowledge might also be disseminated to external people (clients, providers, and privileged partners). Virtual enterprises can distribute knowledge to different organizations with different corporate cultures, which might be important in multinational corporations.

Knowledge distribution can be active (for example, using push technologies such as proactive agents) or passive (using pull

enabling users to retrieve the documents containing a given knowledge statement.

In "Building and Searching an XML-Based Corporate Memory," Auguste Rabarijaona, Rose Dieng, Olivier Corby, and Rajae Ouaddari address the interest in using XML for knowledge management. They propose using XML documents annotated with ontological information to build an organizational memory. They present OSIRIX, a tool that lets users translate a corporate ontology into an annotation document type definition, provided that we represent this ontology in the CommonKADS Conceptual Modeling Language. This tool also lets users retrieve information from XML documents annotated with such ontological information.

Collaborative design over the Internet

Simon Szykman, Ram D. Sriram, Christophe Bochenek, Janusz W. Racz, and Jocelyn Senfaute's article, "Design Repositories: Engineering Design's New Knowledge Base," presents a project for developing a representational infrastructure and a computational framework for creating design repositories. They propose a design-modeling language for representing design artifact knowledge interfaces for creating, editing, and browsing artifact repositories. They offer a design repository tool suite for distributed development of design repositories. This tool suite includes Web-based interfaces enabling access to such design repositories over the Internet by multiple distributed clients using common Web browsers. The suite comprises a Web-based design repository editor and browser. They represent several aspects of the artifact through XML-based schemata.

In "Web-Based Knowledge Management for Distributed Design," Nicholas Caldwell, P. John Clarkson, Paul Rodgers, and Avon P. Huxor describe WebCADET, a Web-based decision-support tool that assists designers—particularly, in a distributed team of designers—in conceptual design evaluation. WebCADET adopts a knowledge server architecture—they situate its knowledge base and its inference engine on a Web server. WebCADET offers design guidance and enables knowledge capture and viewing. It provides a design hierarchy to facilitate the user's navigation through a collection of texts.

technologies).[1] Companies must adapt the distribution to their users' daily work environment.

THIS SPECIAL ISSUE ILLUSTRATES current research in knowledge management. I hope it will stimulate researchers to deepen the links between knowledge management and the Internet and intranets. The worldwide, interdisciplinary research on the semantic Web[5] can find a killer application in knowledge management. Knowledge management will benefit considerably from research, technological progress, and innovative Web and intranet applications, provided that such technology is integrated in the organization to favor knowledge sharing. ■

Acknowledgments

I thank the reviewers for their careful reviews of the 35 articles submitted to this special issue. I also thank the Acacia team members, the Memento working group participants, and the IST Comma project members for their fruitful discussions on knowledge management, the Internet, and intranets.

References

1. G. Van Heijst, R. Van der Spek, and E. Kruizinga, "Organizing Corporate Memories," *Proc. 10th Banff Workshop on Knowledge Acquisition for Knowledge-Based Systems* (KAW '96), B. Gaines and M. Musen, eds., SRDG Publications, Calgary, Canada, 1996.
2. R. Dieng et al., "Methods and Tools for Corporate Knowledge Management," *Int'l J. Human–Computer Studies*, Vol. 51, No. 3, Sept. 1999, pp. 567–598.
3. J. Kingston and A. Macintosh, "Knowledge Management through Multi-Perspective Modelling: Representing and Distributing Organizational Memory," *Proc. Expert Systems 99*, Springer-Verlag, Cambridge, UK, 2000.
4. D. Fensel et al., "Ontobroker: Or How to Enable Intelligent Access to the WWW," *Proc. 11th Workshop on Knowledge Acquisition, Modeling, and Management* (KAW '98), B. Gaines and M. Musen, eds., SRDG Publications, Calgary, Canada, 1998.
5. T. Berners-Lee, D. Connolly, and R.R. Swick, "Web Architecture: Describing and Exchanging Data;" www.w3.org/1999/04/WebData.html (current May 2000).
6. B. Tschaitschian et al., "Internet-Enabled Corporate Knowledge Sharing and Utilization," *Internet-Based Organizational Memory and Knowledge Management*, D. G. Schwartz, M. Divitini, and T. Brasethvik, eds., Idea Group Publishing, Hershey, New York, 2000.
7. D.E. O'Leary. "Using AI in Knowledge Management: Knowledge Bases and Ontologies," *IEEE Intelligent Systems*, Vol. 13, No. 3, May/June 1998, pp. 34–39.
8. J. Euzenat, "Corporate Memory through Cooperative Creation of Knowledge Bases and Hyper-documents," *Proc. 10th Banff Workshop on Knowledge Acquisition for Knowledge-Based Systems* (KAW'96), B. Gaines and M. Musen, eds., SRDG Publications, Calgary, Canada, 1996.
9. A. Farquhar, R. Fikes, and J. Rice, "The Ontolingua Server: A Tool for Collaborative Ontology Construction.," *Int'l J. Human–Computer Studies*, Vol. 46, No. 6, June 1997, pp. 707–727.
10. J. Tennison and N.R. Shadbolt, "APECKS: A Tool to Support Living Ontologies," *Proc. 11th Workshop on Knowledge Acquisition, Modeling, and Management* (KAW '98), B. Gaines and M. Musen, eds., SRDG Publications, Calgary, Canada, 1998.
11. J. Domingue, "Tadzebao and WebOnto: Discussing, Browsing, and Editing Ontologies on the Web" *Proc. 11th Workshop on Knowledge Acquisition, Modeling, and Management* (KAW'98), B. Gaines and M. Musen, eds., SRDG Publications, Calgary, Canada, 1998.
12. B. Gaines and M. Shaw, "Knowledge Acquisition, Modelling and Inference through the World Wide Web," *Int'l J. Human–Computer Studies*, Vol. 46, No. 6, June 1997, pp. 729–759.
13. H. Eriksson, "Expert Systems as Knowledge Servers," *IEEE Expert*, Vol. 11, No. 3, June 1996, pp. 14–19.
14. O. Corby and R. Dieng, " The WebCokace Knowledge Server," *IEEE Internet Computing*, Vol. 3, No. 6, Nov./Dec. 1999, pp. 38–43.
15. "Innovative Applications of the World Wide Web," special issue, B. Gaines, ed., *Int'l J. Human–Computer Studies*, Vol. 46, No. 6, June 1997.
16. "World Wide Web Usability," special issue, S. Buckingham Shum and C. McKnight, eds., *Int'l J. Human–Computer Studies*, Vol. 47, No. 7, July 1997.

Rose Dieng is research director at Inria Sophia Antipolis and scientific manager of the Acacia Project at Inria Sophia Antipolis. Her research interests include knowledge acquisition, knowledge modeling and management, conceptual graphs, and multiagent systems. She has published more than 50 articles in journals, books, or conferences and organized 15 conferences or workshops. She has a PhD in computer science from the University of Paris-Sud, Orsay. Contact her at Inria, Acacia Project, 2004 Route des Lucioles, BP 93, 06902 Sophia-Antipolis Cedex, France; Rose.Dieng@sophia.inria.fr.