KROMOS: ONTOLOGY BASED INFORMATION MANAGEMENT FOR ICT SOCIETIES

Antonio Oliveri, Patrizia Ribino, Salvatore Gaglio, Giuseppe Lo Re
Dipartimento di Ingegneria Informatica (DINFO), Universita' degli Studi di Palermo, Palermo, Italy

Tonio Portuesi, Aurelio La Corte, Francesco Trapani
Sicilia e-Innovazione s.p.a., Italy

Keywords: Knowledge management, Knowledge management systems, Ontology, Software reuse, Rule-base processing.

Abstract: Over the last few years, several projects for the development of innovative systems capable of collecting and sharing information have been carried out, following the increasing companies' interest on a correct knowledge management. ICT companies' managers have realized that knowledge and its management, more than the mere data, constitute fundamental part of their activities. This paper proposes a Knowledge Management System whose main feature is an underlying ontological knowledge representation. This data representation allows the specialization of the reasoning capabilities and the provision of ad hoc behaviors. The system has been designed for the management of projects and processes and has been tested using data coming from projects and processes typical of government ICT companies, providing a Document Management System and an Expert System to share documents and to plan how to best use firms' knowledge.

1 INTRODUCTION

Nowadays most companies work in complex application contexts which create huge amounts of information. Organizations are constantly searching for new solutions to adapt to new conditions in order to survive in increasingly competitive environments. It appears very useful for companies to have awareness of their own information, contained in documents, enterprise processes, acquired experiences and so on. The great amounts of data impose the adoption of new computer-based information systems which enable the storage of structured data and the automation of the information-processing activities of the organization. *Enterprise Knowledge Management*, a new research area, classifies knowledge as a company asset and studies methods and computer technologies to increase its value, reuse and access. Recent studies have drawn attention to problems related to knowledge capitalization and management (Staab et al., 2001),(O'Leary, 1998). During the last two decades ad-hoc frameworks known as *Knowledge Management Systems* (KMS) have been proposed, enabling access and coordination of knowledge assets (Alavi and Leidner, 2005). KMSs represent information systems which manage organizational knowledge with the purpose of increasing the productivity of knowledge operators. This paper is the product of a collaboration between the Computer Engineering Department of Palermo University and the Sicilian local Government ICT society Sicilia e-Innovazione. It proposes an ontology-based knowledge management framework capable of modeling the government offices structure and the ICT company's projects, discovering processes to be automatized and projects to be reused or developed ad hoc. The system we present is a KMS consisting of an expert system for decision support and a document management engine. The main goal of this work is the development of a system for managing data and information that exploits a new approach with a separation between knowledge representation and knowledge management, so that change in the infrastructure of concepts in the knowledge base, done by the expert of the domain, does not influence the inference mechanisms and logical reasoning processes.

The paper is structured as follows: section 2 provides a vision of the state of the art of KMSs and expert systems; in section 3 we introduce essential aspects of KMSs, Expert Systems for decision support and

Oliveri A., Ribino P., Gaglio S., Lo Re G., Portuesi T., La Corte A. and Trapani F. (2009).
KROMOS: ONTOLOGY BASED INFORMATION MANAGEMENT FOR ICT SOCIETIES.
In *Proceedings of the 4th International Conference on Software and Data Technologies*, pages 318-325
DOI: 10.5220/0002254903180325

Ontologies; in section 4 we present a detailed description of the system; section 5 illustrates the Document Management Service of the KMS; section 6 presents our case study. Finally, in section 7, we reach some conclusions.

2 STATE OF THE ART

The challenge of KMS developers is to create a system of tools to collect, organize, and share common data, documents and individuals' expertise. Research has examined different aspects of the problem, studying novel knowledge representation models and involving modern AI techniques to extract new knowledge.
In the last few years, the most important ICT companies have demonstrated an increasing interest in the development of internal knowledge management instruments, many of them based on web corporate portals. A case of a web and ontology-based KMS in a company is the United Nations Food and Agriculture Organization's WAICENT (World Agriculture Information Centre). This platform makes FAO's knowledge about food security and agricultural development widely available to users, with a decision-support systems to improve food security through information use and a document management system allowing people around the world to read and use FAO's documentation (O'Leary, 2008).
L. Razmerita et al. in (Razmerita et al., 2003) proposed a KMS based on the ontological model of the user profile, representing characteristics such as users' preferences, competencies and so on, adopting semantic web techniques. Many other KMSs in literature are designed and customized to satisfy the needs of specific firms. Liping Sui in (Sui, 2005) studied the benefits of a decision support system within the business management. D.J. Harvey and R.Holdsworth in (Harvey and Holdsworth, 2005) observed the advantage of using a KMS in the aerospace industry. An interesting study conducted by Chun, Sohn and Granados in (Chun et al., 2008), shows the use of a Knowledge Management System in an industrial engineering company. Through observation of the company's requests, they identified the features of a KMS which were necessary in order for it to be considered a good investment for the firm.
In this article we introduce a prototypal KMS applied to a real case. Exploiting the advantages of the representation of domain concepts through ontologies, two knowledge management mechanisms have been implemented, one related to efficient document management and the other concerning decisional problems facing efficient government.

3 OVERVIEW

3.1 Knowledge Management Systems

Knowledge Management (KM) consists of a technique that uses Information Technology tools for the management of information, making knowledge explicit and sharing a firm's professional expertises and informative resources. A Knowledge Management Systems, supporting the storage of knowledge, creates the opportunity to make information and knowledge from different sources readily available. KMSs contain both explicit and tacit knowledge. Explicit knowledge, more familiar and easily written down, includes data stored in documents. KMSs can also store tacit knowledge, which is more difficult to express, and includes people's experiences, know-how and expertise. The issue of how to better capitalize and disseminate tacit knowledge is one of the actual priorities in Knowledge Management. To realize such goals, a KMS can make use of different technologies such as:

a. *Document based*: for the creation, administration and sharing of different documents, managing the explicit knowledge.

b. *Ontology/Taxonomy based*: using ontologies and classification for knowledge representation. Knowledge concepts are arranged in hierarchical structures, typically related by relationships. Such methodologies act on both explicit and tacit knowledge.

c. *AI based*: using particular inference engines to resolve peculiar domain problems, they generally manipulates tacit knowledge (g.e. Knowledge-base system).

3.2 Ontologies and Knowledge-based Systems

In computer science the use of the term "*ontology*" means the study of the "*being*", the fundamental categories of which it is composed and the relationships among them to formulate an exhaustive and rigorous conceptual scheme of a particular application domain. Generally it is represented through a hierarchical structure which contains all the noteworthy entities, the existing relationships between them, the rules, the axioms and the specific domain constraints. Knowledge-Based Systems (KBS), a class of AI systems, are able to represent specific domain knowledge

and apply it for solving problems through inference processes. A particular subclass of KBSs is the rule-based expert system, in which knowledge is captured into a set of rules, each encoding a small piece of the expert's skills. Each rule is an "if-then" statement. An expert system emulates the domain expert in the same conditions.

4 KNOWLEDGE MANAGEMENT FOR DECISION SUPPORT: KROMOS

In this paper we present Kromos, a KMS prototype developed through the partnership between the Computer Science Department (Dinfo) of Palermo University and an ICT company of the Sicilian government which deals with the automation of government office processes. Kromos is an ontology-based system of knowledge management with the aim of optimizing business processes for creating and managing ICT projects for government offices. To achieve this, Kromos implements a system of document management and an expert system for decision support.

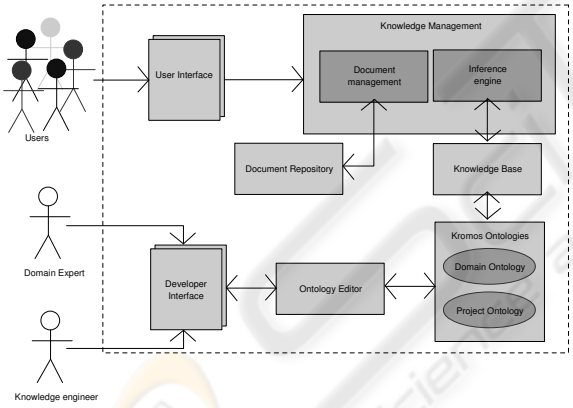

Figure 1: System architecture.

The knowledge representation is based on two ontological domain models, the former reproducing the government offices' structure and the latter modeling the concepts of projects developed by the ICT company. The main idea is the separation between knowledge representation and knowledge reasoning, so that the same infrastructure of rules can be used adopting different knowledge bases. The proposed system was designed with the following features: specific domain knowledge by building a knowledge base, reasoning ability performed by a rule-based expert system, and finally advanced techniques of document and information retrieval. Figure 1 shows the system architecture. The core of the architecture is the Knowledge Management component, while the data level is composed of a repository to store documents and a KB to maintain domain information. All the components are now presented in detail.

4.1 User Interface

The proposed system implements the typical client-server paradigm using JSP and Java servlets (Avedal et al., 2000). With a graphical interface the users can select different application areas. The system is divided in three different macro-areas: PA, Process and Project managements. The first allows the user to manage information about government offices and interacts with domain ontology. The second enables the user to organize new processes in activities that can be potentially automated. The first two areas involve more ontology domain formalization and building processes; the third allows the Project Manager (PM) to access the system functionalities. Using that area the user can create complex queries for the Expert System.

4.2 Knowledge Base

The Knowledge Base of Kromos represents the knowledge container. KB relations and concepts are described using an ontological structure of instances in order to collect and manage data. The rationale behind the Kromos ontology is to provide a minimal but sufficient ontology, suitable for application-domain purpose. Ontologies of the proposed system are performed through Frames (Minsky, 1974) and built using Protégé, a free and open source platform developed by Stanford University, that supports frame-based ontologies according to the Open Knowledge Base Connectivity Protocol (OKBC)(Chaudhri et al., 1998).
In a frame-based model, an ontology is composed by:

- a set of *classes*, hierarchically organized to describe the domain concepts;
- a set of *slots* for the classes, which describes properties and relations between concepts;
- a set of *class instances*, examples of concept with their specific values and properties.

The use of such an ontological model transforms abstract concepts into logical descriptions.

4.2.1 Kromos Ontologies

Modeling knowledge about the government offices world required some assumptions about its structure

and activities, as well as about the nature of the ”observer” expected to use, understand, and rely on the model. Our ontology was designed from scratch for the Kromos KMS and can be considered a collection of two correlated ontologies, a domain ontology and a projects ontology; in order to keep the ontology easy to understand, only a few concepts from the government offices' domain and from computer engineering projects are collected. This results in a simplified description of Projects, Processes and Structure of government offices and a group of details, attributes and relations.

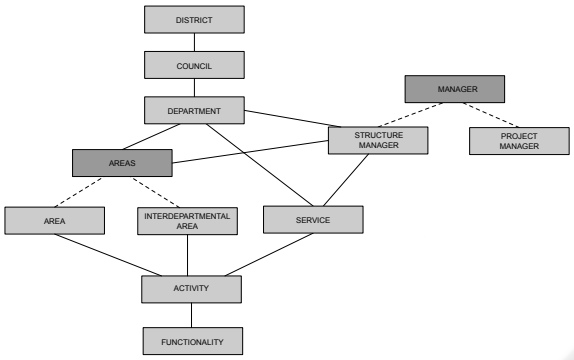

Figure 2: Domain ontology of the Kromos platform.

The **Domain Ontology**, a formal representation of the government offices structure and activities is used to characterize the environment in which the system works, and is organized as a set of concepts and relations allowing deduction of new knowledge; it reproduces the logical architecture of government offices arranged in levels, each depending on the previous in a pyramidal organization (fig. 2).

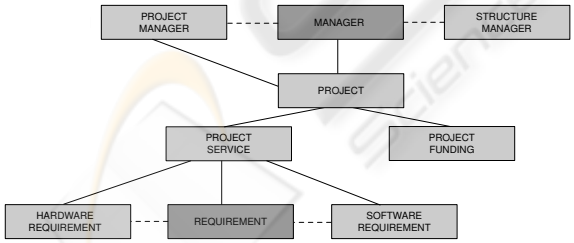

Figure 3: Project ontology of the Kromos platform.

The **Project Ontology** is useful to describe ICT company projects; it maps the structure of the project components containing semi-structured explicit knowledge. During the execution of queries, each component of this ontology is used to get all the elements of the domain in agreement with the query (fig.3).

4.3 Decision Support System

The main goal of an Expert System (ES) for decision support is to assist employees during their activities, finding solutions that usually need the intervention of specifically skilled people. The goal is to incorporate implicit knowledge about the specific field in a computational model. The ES prototype implemented for the Kromos platform is a rule-based system developed in Jess (Java Expert System Shell) (Hill, 2003), that can be used for reasoning in different knowledge base contents, adapting rules to different kinds of domains. It provides rule-based programming suitable for automating an expert system, and is often referred to as an expert system shell. Rather than a procedural paradigm, where a single program has a loop that is activated only once, the declarative paradigm used by Jess matches a rule with a single fact specified as its input and processes that fact as its output. When the program is run, the rules engine will activate one rule for each matching fact. The ES exploits two different kinds of knowledge: declarative facts, captured by the ontological model, and procedural facts, expressed using rules defined by an expert.

Use cases: Use cases in which our ES can offer support are decisional processes such as:

a) *Project planning process* - During planning of company projects there are many different constraints to be considered in order to improve enterprise yields and avoid wasting resources. Planning is a process for the definition of a future goal, the activities to exploit in order to reach that objective, and all the resources to be used to complete these activities. The planning process has to identify business components directly connected to the real progress of business activities, measuring their impact and connected benefits, and to analyze the investment policy, producing a Business Plan, as figure 4 shows.

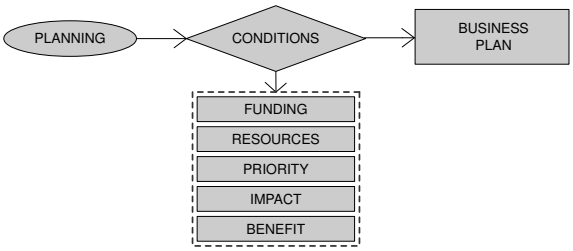

Figure 4: Use case a: Project planning process.

b) *Project management process* - Supervision of the state and progress of projects. This process guides the business management to the attainment of previously planned objectives, showing the differ-

ences between them and the results obtained, so that managers can decide and actuate appropriate corrective actions. For instance, figure 5 illustrates that process.

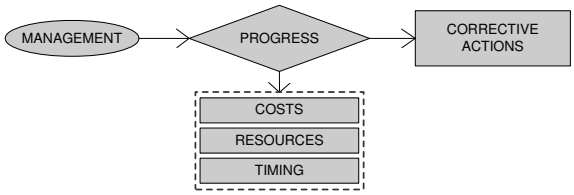

Figure 5: Use case b: Project management process.

c) *Evaluation of previous projects' functionalities reuse* - Analysis of government requirements and already developed projects to find reusable components. The organization we are analyzing produces a great amount of ICT products to automatize district processes. A process is a set of interrelated activities, grouped in phases. Therefore each project is composed by a set of components, each supporting a singular phase of the entire process. Different district processes could have certain phases in common, so that the organization could choose to reuse some components taken from other projects during automation activities, in order to reduce developmental costs.

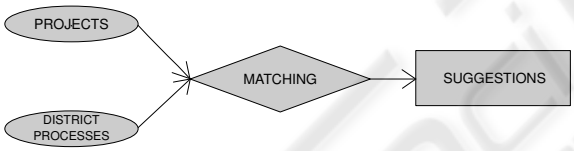

Figure 6: Use case c: Evaluation of previous projects' functionalities for reuse.

For instance, figure 6 shows the ES used to identify which project components could be reused to automate district processes phases.

4.3.1 Examples of Rules

The code portion reported below shows a typical ES behavior in the third use case. In this scenario the expert system uses the declarative part of the knowledge that expresses the active part of procedural knowledge emulating the behavior of a human expert.

```
(defrule search_department ?instance
<-(object (is-a DEPARTMENT)
(name ?n&:(call ?*ric* equals
(slot-get"+ Dep+"name)
(lowcase ?n))))=>(bind $?area
(slot-get (instance-name ?instance)
comprehend_areas
(foreach ?j $?area
(if  (call ?*ric* different
(slot-get (instance-name ?j) name) empty)
then...
```

5 KROMOS DOCUMENT MANAGEMENT SERVICE

The Document Management Service (DMS) provided by Kromos is a module capable of pre-processing documents, retrieving data, indexing texts and search engine managing. In ICT companies the volume of documents produced during working activities grows rapidly, collecting them in traditional forms becomes almost impractical and also searching them without automatic search engines is a great waste of time. Documents contain most of the information about projects, functionalities, people involved and so on. The Kromos platform uses Apache Lucene, an Information Retrieval (IR) engine adapted for the insertion, indexing and retrieval of documents in different formats. Lucene is an open-source, high performance, scalable, full-featured text search engine and information retrieval library written in Java and suitable for any application requiring a full text search (Gospodnetic and Hatcher, 2005), through which any piece of data convertible to a textual format can be indexed and made searchable (Pirro and Talia, 2007). The main functionalities of the Kromos Document Management Service are: pre-processing of documents and their content to obtain a text representation without any lexical or sematic redundance; document indexing to store information about files in an ordered structure to use for the search phase; searching for documents using keywords, calculating the degree of satisfaction of the requirements expressed in the query. These requirements are fulfilled by three Kromos modules: *pre-processing*, *indexing* and *searching*.

5.1 Pre-processing Module

Pre-processing is a necessary procedure in document management, through which data and information stored in documents in a specific format can be elicited by analyzing and tokenizing content. Organizations generally create and use a great amount of documents that can be stored in different kinds of formats like text files (.txt), document files (.doc, .pdf), web pages (.xml, .html) (Zhou and Xie, 2007). The analysis of heterogeneous format contents, the removal of meaningless terms and the maintenance of information useful to retrieve and recover docu-

ments will depend on the DMS. The proposed DMS works in different steps in order to pre-process and organize documents. When an organization's members insert documents to be collected by the KMS, a two-step parsing then occurs: *content elicitation* and *content tokenization*. Content elicitation to withdraw textual content from different kinds of files eliminates irrelevant information, such as typesetting format, and transforms content into a character data stream. Content tokenization breaks the content into words and sentences and transforms the data stream into a set of terms for the subsequent content parsing procedure.

5.2 Indexing Module

Apache Lucene was adopted for indexing and document storage, the search interface for querying index and the reading interface to read texts and documents. The fundamental concepts in Apache Lucene are index, document, field and term (Bennett, 2004) (fig. 7).

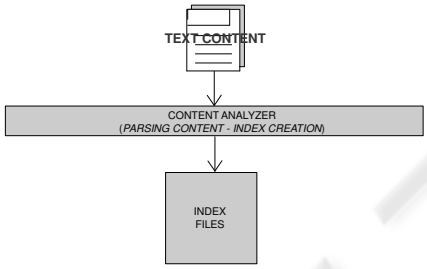

Figure 7: Index creation in document management system.

An index contains a sequence of documents, which are a sequence of fields. Each field gets tokenized and generates pairs of field name and text tokens called terms. The index stores input in a data structure called inverted index, making efficient use of disk space while allowing quick keyword lookups. Kromos provides a user interface for document insertion, giving users the opportunity to collect them in a unique repository.

5.3 Searching Module

After the initial collection of documents and the creation of an index, users need to retrieve documents from the remote file system, seeking them using a search interface. The user interface is also a web-based interface, so that user works as if the interface were a common web search engine, typing keywords, title, project name and so on. The system searches in the indexed documents and retrieves the relevant documents, as figure 8 shows.

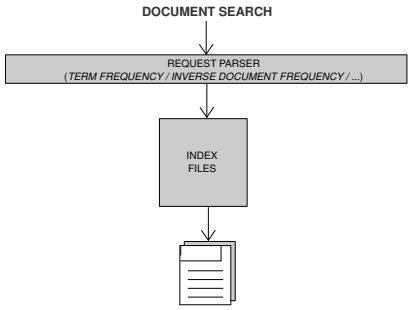

Figure 8: Document search in document management system.

The result is a list of documents, each linked to the related file; when opened, the system displays a reading interface in a *What You See Is What You Get* view. In Kromos documents are retrieved on the basis of their content as if they were composed by atomic text entities, a set of different concepts considered here by calculating their frequency in the document, disregarding the document's overall structure.

6 CASE STUDY

As previously mentioned, Sicilia e-Innovazione is the research and development company of Sicilian local government in the ICT field, and represents one of the most interesting setting in order to acquire experiences for the development of a Knowledge Management System. This is due to several reasons such as the novelty of the company mission, the heterogeneity of data sources, the difficulty of adopting traditional tools, and the involvement of different users. Moreover, previously the start of the project, the problem of data and knowledge management had received small attention. In that period, only few documents on business processes and organization of government offices were produced, while a huge amount of technical documents were produced and stored without any management organization. All these reasons make this company an ideal scenario where experimenting a process of knowledge engineering, in order to promote the adoption of a knowledge management system as tool for sharing the already acquired knowledge, and for the production of the new one.

The Previous Situation. Sicilia e-Innovazione was created in 2006 and since its establishment carried out a lot of projects, mostly for the automatization of all for government office processes. Each project is described and documented by several technical documents that represent a valuable source of precious knowledge. Each document contains descriptions of

the systems, in its current and future forms, financing sources, responsibility information, technical and developing choices, and the processes to be automated. All this knowledge is precious and could be used in several different ways, for instance, to estimate costs and necessary resources for new projects, to find similarities in different contexts, to increase the code reusability. When the project started, the company was organized as a collection of atomic groups, each working independently by the others; different teams replied the same development activities, so reuse was minimal, there was no effort to organize and share experiences and knowledge acquired by previous projects, and any technical tool (such as search engines) to search and compare different projects. This situation produced knowledge islands not shared among product groups and program teams, with generational gaps between seasoned and newly hired employees, limited ability to learn from existing knowledge and no unified vision.

The Application of Kromos. The experience carried out during the process of development in the bosom of Sicilia e-Innovazione gave us the chance to define an ontological representation of the specific domain, and to test our general-purpose Knowledge Management prototypal system in a real applicative environment.
Our working activity in the firm consisted of two different phases: the first one, after the acquisition of specific knowledge through the observation of business activities and processes, allowed us to build an ontological model for the domain representation; the latter was centered on the fitting of Kromos reasoner and document management services to the company's needs. In the meanwhile some new requirements emerged and that permitted us to refine our system with new features in order to improve Knowledge accessibility to different skilled users. This experience provided all the information to comprehend the importance of a correct and efficient knowledge resources, people and organization management; to estimate the efficacy of adopted technical solutions, their advantages and disadvantages; and to evaluate needs satisfaction degree using these solutions. The document management and information acquisition mechanisms, such as those developed for this prototype, proved to be necessary in a dynamic environment like the one we observed.

Benefits using the KMS Kromos. The adoption of Kromos for the data management evidenced the advantages of a correct knowledge management in every organization, emphasizing the importance of sharing and reusing information with different skilled workers. The use of a modifiable ontology below the expert system gave flexibility to the platform, so that its use is also an activity which continually refines the knowledge base. During use of the system, more information was collected and ontology structure changed and grew, but the functionalities of the expert system did not require any modification. The web based interface, hiding the complexity of the system's functionalities, gave the platform a characteristic of usability, so that the collection of data and documents and the sharing of information seemed to be incremented.

7 CONCLUSIONS

The key thrust of this article has been to analyze the benefits of knowledge-based systems for knowledge management and the definition of expert system rules which can adapt their results to changes in ontology. Reuse and information sharing are essentially Knowledge Management problems. From this point of view Kromos provides: an improvement of knowledge sharing between employees because it accelerates the transfer of knowledge; an improvement on the knowledge retrieval through document management system; an improvement in knowledge reuse through an expert system able to provide a support in the adaptation of old solution to new problems. The use of such a system in a real case, an ICT Sicilian Government company, gave the opportunity to measure the increase in information sharing and reuse, seconding the advice obtained by the elaboration of the expert system.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the contribution of Marco Sicilia and Mirella Marrone to the support for Kromos graphics.

REFERENCES

Alavi, M. and Leidner, D. (2005). Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *Knowledge Management*.

Avedal, K., Halberstadt, A., Ayers, D., Briggs, T., Burnham, C., Haynes, R., and Zeiger, S. (2000). *Professional JSP*. Wrox Press Ltd. Birmingham, UK, UK.

Bennett, M. (2004). Contrasting relational and full-text engines. NIE (New Idea Engineering) Enterprise Search Newsletter 2 (9), article 1.

Chaudhri, V., Farquhar, A., Fikes, R., Karp, P., and Rice, J. (1998). OKBC: a programmatic foundation for knowledge base interoperability. In *Proceedings of the fifteenth national/tenth conference on Artificial intelligence/Innovative applications of artificial intelligence table of contents*, pages 600–607. American Association for Artificial Intelligence Menlo Park, CA, USA.

Chun, M., Sohn, K., Arling, P., and Granados, N. (2008). Systems Theory and Knowledge Management Systems: The Case of Pratt-Whitney Rocketdyne. In *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*, pages 336–336.

Gospodnetic, O. and Hatcher, E. (2005). *Lucene in Action*. Manning.

Harvey, D. and Holdsworth, R. (2005). Knowledge management in the aerospace industry. In *Professional Communication Conference, 2005. IPCC 2005. Proceedings. International*, pages 237–243.

Hill, E. (2003). *Jess in Action: Java Rule-Based Systems*. Manning Publications Co. Greenwich, CT, USA.

Minsky, M. (1974). A Framework for Representing Knowledge.

Nonaka, I. (2005). A Dynamic Theory of Organizational Knowledge Creation. *Knowledge Management: Critical Perspectives on Business and Management*, 5(1):14–37.

O'Leary, D. (1998). Enterprise Knowledge Management. *Computer*, pages 54–61.

O'Leary, D. (2008). A multilingual knowledge management system: A case study of FAO and WAICENT. *Decision Support Systems*, 45(3):641–661.

Pirro, G. and Talia, D. (2007). An approach to Ontology Mapping based on the Lucene search engine library. In *Proceedings of the 18th International Conference on Database and Expert Systems Applications (DEXA 2007)-Volume 00*, pages 407–411. IEEE Computer Society Washington, DC, USA.

Razmerita, L., Angehrn, A., and Maedche, A. (2003). Ontology-Based User Modeling for Knowledge Management Systems. *Lecture notes in Computer Science*, pages 213–217.

Staab, S., Studer, R., Schnurr, H., and Sure, Y. (2001). Knowledge Processes and Ontologies. *IEEE Intelligent Systems*, pages 26–34.

Sui, L. (2005). Decision support systems based on knowledge management. In *Services Systems and Services Management, 2005. Proceedings of ICSSSM'05. 2005 International Conference on*, volume 2.

Sumiya, S. and Saito, T. (1992). Development of a multimedia document management system for cooperative work environment. *Computer Software and Applications Conference, 1992. COMPSAC '92. Proceedings., Sixteenth Annual International*, pages 346–355.

Takeuchi, H. and Nonaka, I. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford University Press New York.

Zhou, D. and Xie, K. (2007). Lucene Search Engine. *Jisuanji Gongcheng/ Computer Engineering*, 33(18).