A Prototypal Environment for Collaborative Work within a Research Organization

M. Ettorre[1], L. Pontieri[3], M. Ruffolo[1,3], P. Rullo[1,2], D. Saccà[1,2,3]

[1]*Exeura s.r.l. Via P. Bucci 87030 Rende (CS), Italy E-Mail: {ettorre,ruffolo}@exeura.it*

[2]*Università della Calabria – Dip. DEIS. – Dip. Di Matematica, Via P. Bucci, 87030 Rende (CS), Italy E-Mail: {rullo,sacca}@unical.it*

[3]*ICAR-CNR, Via P. Bucci, 87030 Rende (CS), Italy E-Mail: pontieri@icar.cnr.it*

Abstract

A research organization is a complex system consisting of various kinds of components and actors with a high level of cooperation. Collaborative work takes place through both formalized groups (such as sections, internal committees, research groups, project groups) and informal groups (i.e., communities of practice) that may arise around common problems, interests and objectives (e.g., a research topic, the production of a publication, the use of a technology, etc.). It may happen that groups are geographically distributed over different locations. From this the need of technologies and tools addressing the information sharing needs within research organizations.

This paper presents an experimental knowledge management tool, the ICAR Knowledge Portal, running at ICAR-CNR, a research institute of the Italian National Council of Research. The task of this tool is basically to provide a web-based environment for knowledge creation, sharing and access. In fact, with the ICAR Knowledge Portal, users can quickly set up their own sub-portals, create and organize knowledge items (such as news, papers, links, announcements, projects, etc.), share them within workgroups, search for a specific document or browse through a set of related documents (ontology-driven browsing).

1. Introduction

One of the main areas of ICT is nowadays represented by Knowledge Management (KM). The aim of this technology is to support the management and the application of knowledge within organizations - both companies and public administrations. The socio-economic and technological transformations of the last few years have indeed turned organizations into "knowledge intensive" structures that, in order to be competitive in the age of globalization, need to constantly increase their know-how. The new organization paradigm is that of "learning organization", i.e., a structure able to acquire, organize and transfer knowledge, and to modify its behaviour according to its knowledge about the reality. From this the need, for a modern company, to be equipped with advanced software environments giving the ability to manage knowledge.

The task of a knowledge management environment is to act as an infrastructure to support the knowledge lifecycle within an organization, thus providing all the necessary features for acquiring, modelling, sharing, retrieving and publishing knowledge [1,13].

Knowledge acquisition is concerned with the recovering of information that is spread around - both, inside and outside the organization [7]; thus, it relies on techniques designed for gathering and integrating knowledge from multiple sources, making implicit knowledge explicit, extracting knowledge from unstructured or semi-structured data, and so on. Qualifying technologies are basically those coming from the fields of Knowledge Discovery (data and text mining), Information Retrieval and Natural Language Processing.

Knowledge modelling deals with the formal representation of the acquired knowledge. To this end, ontologies can be used as suitable tools for structuring the knowledge about the application domain in terms of axioms, concepts and relationships. Broadly speaking, an ontology can be regarded as an agreed conceptualization of some area of content. There are several methodologies [2,4,6] and formalisms [3,8,9,10,11,14] for this purpose, many combining classical database abstraction mechanisms with representation techniques coming from Artificial Intelligence.

Knowledge sharing has to do with collaborative work inside an organization. A key concept in this context is that of communities of practice, i.e., "informal aggregations of people who share work practices and common experiences". Within an organization, a knowledge management approach can address the collaborative-work challenge, by enabling information sharing for both formal and informal groups.

Knowledge access gives users the ability to easily retrieve knowledge items [12]. In principle, two basic approaches are possible: first, a navigational approach, for instance, based on an ontology-driven classification of contents - the ontology structure directs the user to locate documents related to a given subject; second, a retrieval approach based on a declarative specification of the properties that are relevant for a given concept - technologies for this purpose are essentially based on full-text search techniques or on a structured representation of contents (metadata) allowing for classical retrieval techniques.

This paper presents the experience of design and implementation of a knowledge management environment carried out at "Istituto per il CAlcolo e le Reti ad alte prestazioni", a research institute of the Italian National Council of Research (ICAR-CNR for short). ICAR-CNR operates in the area of Information Technology and has seats in three different regions of Southern Italy. It is a very complex organization where collaborative work takes place within formal groups (committees, research groups, project groups, etc.) as well as communities of practices, often spread across different seats. From this, the need to provide researchers with means to organize and share contents throughout all the organization.

The ICAR Server is the infrastructure designed and implemented to create the ICAR Knowledge Portal. It provides the basic services for knowledge organization, sharing and access. The ICAR Knowledge Portal is a web-based application, built on top of the ICAR Server, supporting researchers in the management of the knowledge produced inside the ICAR-CNR institute. It is built around a set of ontologies which provide a map of the knowledge about the ICAR-CNR application domain. With the ICAR Knowledge Portal, users can easily create their own sub-portal, create and organize knowledge items (news, papers, links, announcements, projects, etc.), share them within workgroups, or make them public over the Internet. Further, they can quickly access documents, either by browsing through information by concepts (ontology-driven browsing) or searching for information through documents properties.

The paper is organized as follows. In Section 2 we give a presentation of ICAR-CNR in terms of organizational structure, interest areas and knowledge management problems. In Section 3 we outline the basic features of the ICAR Knowledge Server. In Section 4 we give an overview of the ICAR Knowledge Portal and, finally, in Section 5 we provide our conclusions and further work.

2. The ICAR-CNR Institute

The "Istituto per il Calcolo e le Reti ad Alte Prestazioni" is a research institute of the Italian National Council of Research (ICAR-CNR, for short) with sections in Cosenza (the head office), Naples and Palermo. The mission of the institute is to carry out research activities in the field of Computer Science and Information Technology, cooperating with the academic world and with other private and public research organizations. Some of the prominent research topics the Institute is engaged in are: Database and Knowledge Base Systems, Models and Tools for Parallel Processing, Representation and Control of Complex Systems, Knowledge Discovery in Databases and Data Mining, Artificial Intelligence, Multimedia Systems, Mathematical Computation, Artificial Vision.

In addition to its research activities, the institute is interested in the pre-competitive development and technological transfer of research results, and carries out educational and training activities, through scholarships and research fellowships, advanced after-university specialization courses, and non-university higher education activities.

The scientific and administrative responsibility of the Institute is assigned to a Director, while for each section there is a Section Responsible. The management roles take advantage of the support of the following committees: the Scientific Council, the Institute Committee, and the Section Committees (one for each section).

People at ICAR-CNR may work either in the administrative or in the research area in several different positions. Furthermore, scholarship students, fellowship students and part-time people collaborate with the institute's staff.

Researchers usually work on projects, proposed and developed in conjunction with public and private organizations; further, they can be involved in other activities, such as teaching at university. For each scientific theme of interest, there is a research area with one responsible; likewise, for each research project, there is a project team. A researcher can be engaged in more project groups and can be tied to different research themes.

As it is natural within a research organization, collaborative work takes place through a number of

communities of practices. Indeed, in addition to formalized groups (sections, internal committees, research groups, project groups) many informal groups may arise, often spontaneously, around common problems, interests and objectives (e.g., a research topic, the production of a publication, the use of a technology, bureaucratic and administrative issues, etc.) and may evolve in the time. We note that such groups might be spread over different locations of ICAR-CNR. In summary, the institute behaves as a complex system, consisting of various kinds of components and actors with a high level of cooperation.

3. ICAR Knowledge Server Overview

The ICAR Server is the infrastructure designed and implemented to create the ICAR Knowledge Portal. Although conceived for a specific purpose, its design and implementation features make it flexible enough to be considered as a general tool for the creation of web-based environments for knowledge management within research institutions.

The ICAR Knowledge Server provides the basic services for ontology specification, user management and, further, knowledge creation, organization, sharing and access.

Ontology Definition. The ICAR Server provides an interface for the definition of ontologies. It is based on a simple graphical language (UML-like) supporting the specification of concepts and relationships. Concepts may represent actors, actions or objects and may have descriptive attributes. Relations are either taxonomic (IS_A and PART_OF) or non-taxonomic.

User Management. There are four types of users:

- single users, also called *knowledge workers*
- *project groups* consisting of knowledge workers formally involved in the same project,
- *communities of practice* (also called groups of interest) consisting of knowledge workers with common scientific interests,
- *confidential groups*, i.e., unofficial, "private" communication channels.

Each user, except the latter, has its own sub-portal. The ICAR Knowledge Server uses *functions* and *rights* to identify responsibilities and roles of knowledge workers. There is a predefined function, the *portal responsible*, who can assign the sub-portal responsible and the sub-portal publisher functions to other knowledge workers. A sub-portal responsible can designate one or more *publishers* for his/her own sub-portal; each publisher is in charge of validating requests for knowledge item publishing through the sub-portal under his control. Furthermore, the portal responsible can grant rights to knowledge workers according to their position in the organization. The *community creation* and *project creation* rights give permission to create communities of practice and project groups, respectively, while the *My-Portal visibility* right allows knowledge workers to make their own personal portal visible on the web. Notice that no grant is required to create a confidential group.

Knowledge Creation and Organization. With the ICAR Knowledge Server, (authorized) users can easily and quickly create their own sub-portals; these are basically of two types: personal sub-portals (MyPortal) and team sub-portals which, in turn, can be classified as either

- project sub-portals, supporting collaborative work within project groups,
- community sub-portals, supporting activities within communities of practice.

Within a sub-portal, a user can create and store different kinds of knowledge items, such as links, news, documents, announcements, software, etc. Before storing a knowledge item (let us call it KI), its creator is asked to annotate it by a number of metadata, i.e., structured information pertaining to KI such as authors' name, keywords, date of creation, etc.; after that, he/she may proceed to perform one of the following actions:

- associate KI with one or more concepts of the underlying ontologies,
- link KI to other knowledge items,
- share KI with other users,
- publish KI over the Internet.

Knowledge Sharing and Publishing. Sharing the knowledge item KI is the process whereby KI is made available to other users of the intranet (i.e., knowledge workers, project groups, communities of practice or confidential groups). A grant on KI defines the rights a user holds on KI; it is either of type OWNER = {read, write, eliminate, update grant, remove grant, assign grant} or of type ACCESS, i.e., any subset of {read, write}. That is, an ACCESS grant on KI enables a knowledge worker to a read/write access mode (on KI), while a grant of type OWNER, besides the read/write permissions, gives him\her the following privileges:

- eliminating KI,
- updating or removing ACCESS grants on KI
- assigning new grants (of both OWNER and ACCESS type) on KI to other users.

We notice that, in order a knowledge worker assigns a grant on KI to a confidential group, he/she must be a member of that group. This is because a confidential group is conceived as a sort of private communication area, not visible to external users. From the viewpoint of the sharing process, a confidential group actually acts as

a shortcut or a distribution list among a predefined set of knowledge workers.

Now, the granting policy supported by the ICAR Server is as follows. The author of KI is assigned with an OWNER grant on KI (thus, we say that he/she is an owner of KI, actually, the first owner). Thus, he may set or modify the access rights on KI of other users; in particular, he can (transitively) grant users the OWNER privileges. As a consequence, the knowledge item KI may have many different owners at the same time. Hence, it may happen that a user U, who already holds a grant on KI, say G1, receives a new grant, say G2, from some owner of KI; now, the grant G assigned by the system to U on KI is $G = G1 \bigcup G2$. Thus, if for instance U has the ACCESS grant {read} on KI and is privileged with the new ACCESS grant {write}, then his/her grant on KI becomes {read, write}. Likewise, if U owns any ACCESS grant on KI and then is qualified with an OWNER grant on KI, he/she is promoted to be owner of KI (indeed, for any ACCESS grant G, $G \subset$ OWNER holds).

Publishing the knowledge item KI, which is stored on a given sub-portal, is the formal process whereby KI is made visible on the web (we say that KI is public). To this end, an owner of KI must require the authorization to one publisher of that sub-portal. The ICAR Knowledge Server supports mechanisms whereby the request is automatically displayed on the "desk" of all publishers of that sub-portal. The request is accepted if there is one publisher who validates it.

Knowledge Access. The ICAR Knowledge Server provides features that help users to easily access portal contents. In particular, there are three ways of searching contents:

- browsing by categories, where ontology structures are exploited to direct users to the searched knowledge items through the topics they are related to,
- direct search through metadata, where classical searching techniques on structured data are applied,
- full-text search, used to implement queries which require a deep analysis of document contents (e.g., looking for a phrase, performing proximity searches, etc.).

In the latter case, efficiency can be improved by preliminarly narrowing the search scope, for instance, by applying selections on metadata or by focusing on a specific topic through ontology browsing.

4. ICAR Knowledge Portal Overview

The ICAR Knowledge Portal is a web-based application, built on top of the ICAR Server, allowing users to create, share, publish and access knowledge produced inside the ICAR-CNR institute.

Broadly speaking, the design of the ICAR Knowledge Portal has been focused around two main activities:

1. recognizing the contents and services that are to be provided by the system in order to meet the user needs,
2. formalizing the knowledge of the application domain in terms of ontologies.

To this end, many different aspects of the application domain have been analysed, such as:

- organizational structure in terms of areas (administration, direction, logistics and so on), responsibilities, staff organization, human relationships, professional roles, etc.
- sources and kinds of explicit knowledge (usually contained in electronic texts, databases, libraries, etc.),
- sources and kinds of implicit (tacit) knowledge (owned by people),
- presence of communities of practice, their structure and topics of interest,
- typology of users, their interests, needs and expectations.

Given the set of desiderata and requirements coming out from the above analysis, the implementation of the ICAR Knowledge Portal essentially consisted of the following steps:

- specification of the ontologies that are required by the system,
- creation of an initial set of sub-portals (the institutional ones),
- creation of an initial set of users with the respective profiles (i.e., functions and rights).

The ontologies underlying the ICAR Knowledge Portal are taxonomies describing different aspects related to knowledge management at ICAR-CNR; among those, the Research Ontology (RO) and the Knowledge Item Ontology (KIO).

The taxonomic hierarchy representing the RO is depicted in figure 1. As we can see, there are three research themes, namely, Parallel and Distributed Computing, Protocol and Technologies for Advanced Applications on Networks and Intelligent Systems; each theme is in turn subdivided into a number of sub-themes. The RO is common to all branches of the ICAR Institute. The RO is the key for content-based (ontology-driven) access to knowledge items, as it allows to organize all knowledge items produced in different locations by different actors.

The KIO provides the ICAR Server with all the necessary information for knowledge item management. The taxonomy shown in figure 2 is based on two kinds of

hierarchical relations: IS_A (continuous lines) and PART_OF (broken lines). Further, there is a non-hierarchical relation (IS_RELATED_TO) expressing that knowledge items can be linked to each other. For an instance, an electronic document can be linked to other conceptually related knowledge items. As an example of PART_OF relation, consider the knowledge item Curriculum: as we can see, it is composed of a number of other knowledge items, such as Biographic Notes, Teaching Activities, Projects Description, Publications, etc.; an interesting consequence of such a definition is that it makes possible the automatic composition of a curriculum starting from its components. Finally, an example of IS_A relation is given by the representation of the knowledge item Pubblication that is expressed as a generalization of ICAR Publications (i.e., papers having at least one ICAR researcher among the authors), non-ICAR Publication and Technical Report.

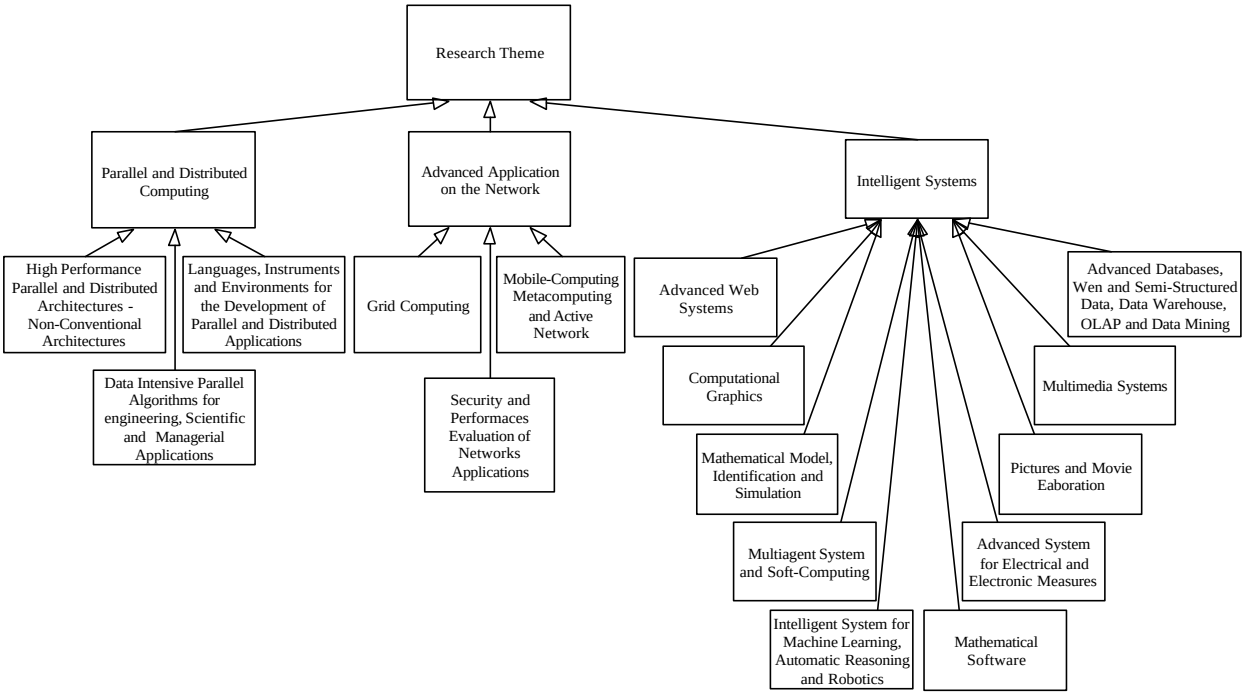

Figure 1: The Research Taxonomy

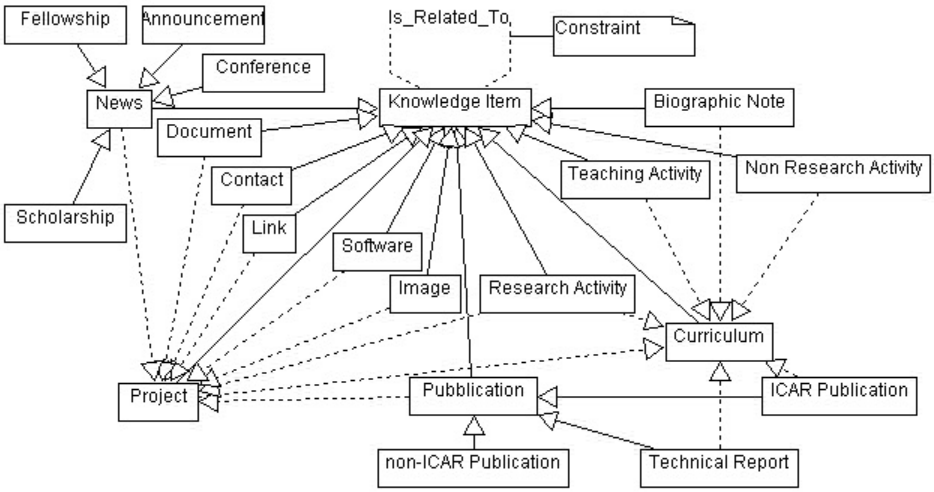

Figure 2: The Knowledge Item Ontology

Once ontologies have been provided to the Knowledge Server, the construction of the following institutional sub-portals has been carried out:

- the institute sub-portal which provides institutional information concerning the institute as a whole (http://kms.icar.cnr.it),
- the branch sub-portals, one for each seat of the institute, which offer local information,
- the committees sub-portals to support knowledge circulation and sharing within internal committees (the Scientific Council, the Institute Committee).

The last implementation step consisted in the creation of an initial set of user functions and rights, to provide the basis for team sub-portals creation, thus making the system evolving in the time. To this end, the responsible of the ICAR Knowledge Portal (who is a predefined function) assigned the publisher functions for the institute sub-portal as well as the branch sub-portal responsibilities. Further, he/she granted the community creation and project creation rights to give permission to construct communities of practice and project groups, respectively.

Since its creation, the ICAR Knowledge Portal has rapidly grown. Today, the system supports several team sub-portals (visible within the intranet) and confidential groups, thus acting as a valuable tool inside the scientific community of ICAR-CNR. To give an example, let us consider the case of a project on High-Performance Computational Grids, where ICAR-CNR is involved with all three of its branches (thus, we are in presence of a workgroup distributed over different geographic locations). The creation of a sub-portal for this project has strongly simplified the interactions among the researchers, allowing them to share knowledge items inherent the research activities tied to the project, such as papers, news, links about relevant related documents or events, etc. To this end, the sub-portal provides facilities such as a bibliographic area where people share publications on Computational Grids and other topics of interest for the project, and a resource area where researchers share documents produced within the project (e.g., technical reports, papers, images, software). An interesting aspect of the sub-portal (actually, of all team sub-portals) is that, whenever a knowledge item is added, it is automatically associated to the (sub-)theme of the Research Ontology to which the project belongs ("Mobile-Computing, Meta-Computing and Active Network", in our case). This corresponds to a sort of automatic categorization which makes it possible a content-based access to all knowledge items produced within the project (ontology-driven retrieval).

5. Conclusion and Future Work

This paper has presented the experience of creation and utilization of a web-based environment for knowledge management within a research institution. We have first shown the basic services provided by the ICAR Server, the infrastructure created and implemented to create the ICAR Knowledge Portal, and then we have seen how this portal has been customized to become a real valuable support for collaborative work among researchers at ICAR-CNR. Throughout the paper, we have often underlined the central role played by ontologies as they provide the formal description of concepts, relationships and constraints that characterize the application domain. Ontologies are exploited for knowledge items categorization and access.

In its current version, the ICAR Server essentially provides services for knowledge item creation, sharing, publishing and access. In order to go beyond the scope of an experimental system, a number of new functionalities should be added, such as automatic retrieval (for instance from the Web) and classification of knowledge items, concurrent creation of electronic documents, content presentation (i.e., personalization and adaptation), intelligent workflow, and so on. To this end, a number of research challenges should be addressed. Ongoing work is mainly focused on (semi-)automatic retrieval and classification and intelligent workflow.

6. References

[1] S. Barnes, "Knowledge Management Systems. Theory and Practices", Thomson Learning, Oxford, 2002.

[2] J. Bouaud, B. Bachimont, J. Charlet and P. Zweigembaum, "Methodological principles for structuring an ontology" *In ACM Press*, editor, Proc. of IJCAI95 Workshop: Basic Ontological Issues in Knowledge Sharing, 1995.

[3] DAML+OIL. Specification. In www.daml.org/2001/03/daml+oil-index.htm, 2001.

[4] M. Fernandez, A. Gomez-Perez and N. Juristo "Methontology: from Ontological art Towards Ontological Engineering", *In Proc. of the Spring Symposium Series on Ontological Engineering* (AAAI97), 1997.

[5] D. Fensel, "Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce", Springer, Berlin, 2001.

[6] N. Guarino and C. Welty, "A formal ontology of properties" In R. Dieng and O. Corby, editors, Knowledge Engineering and Knowledge Management:

Methods, Models and Tools. International Conference EKAW'2000, pp. 97-112. Springer-Verlag, 2000.
[7] Hayes-Roth, F., Waterman, D.A. and Lenat, D.B. Building, "*Expert Systems Reading Mass*", Addison-Wesley, 1983.
[8] Hendler, J., and McGiuness, D., "The DARPA Agent Markup Language", IEEE Intelligent Systems 15 (6), 2000, 72-73.
[9] S. Luke and J. Heflin, "SHOE 1.01. Propose specification", www.cs.umd.edu/projects/plus/SHOE/spec1.01.html, 2000.
[10] N. Noy, "Tutorial on Ontology Engineering", In International Semantic Web Working Symposium (SWWS'2001). www.Semanticweb.org/SWWS/program/tutorials/tutorial1/2001
[11] RDFS, Resource description framework schema specification. In www.w3.org/TR/rdf-schema/, 2002.
[12] Spark-Jones97, K., and Willet, P., "Readings in Information Retrieval, Mountain View" Morgan Kaufmann, 1997.
[13] A. Tiwana, "the Knowledge Management Toolkit", Prentice All, USA, 2002.
[14] van Harmelen, F., and Horrocks, I., "FAQs on OIL: the Ontology Inference Layer", IEEE Intelligent Systems 15(6), 2000, 69-72.