P2P Interactions for the Support of Knowledge Sharing in Networked Enterprises

Roberto RATTI[1], Albert BOKMA[2], Kevin GINTY[2], Dimitrios TEKTONIDIS[3], Adamantios KOUMPIS[3]
[1]*TXT e-solutions, Via Frigia 27, Milan, 20126, Italy*
Tel: +39 02 25771230, Fax: + 39 02 2578994, Email: roberto.ratti@txt.it
[2]*University of Sunderland, St. Peter's Campus – SCAT- Room DM401, Sunderland, SR6 0DD., United Kingdom, Tel: +44 191 515 3233, Fax: + 44 191 515 3236, Email: albert.bokma@sunderland.ac.uk*
[3]*ALTEC Research Programmes Division, M. Kalou 6, GR 546 29 Thessaloniki GREECE*
Tel: +30 2310 595 659, Fax: +30 2310 595 630, Email: [akou][dte]@altec.gr

Abstract: The paper presents results from NOPIK, a current IST project, which is in its final stages. A number of tools to support team-based problem solving activities have been developed and evaluated with the project users and with overall positive results. The toolset also sports a distributed architecture where separate installations and their content can be synchronised with a central server. This in principle provides an architecture enabling collaboration and sharing in a relatively static environment, but in an open world where ad hoc connectivity is required a peer to peer architecture is needed and which is proposed and discussed in this paper. Our research has shown that organisations are usually sensitive to letting documents be stored off their premises and thus solutions such as P2P offer the potential to overcome these barriers and provide an ad hoc connectivity otherwise not easily achievable.

1. Introduction

The aim of our work is to develop tools to create and manage open personal knowledge spaces that can be connected to provide shared network resources. Information is growing at an alarming rate and exists in a variety of formats and locations. This requires a variety of tools to access and navigate it. Users typically develop in their daily interaction a considerable amount of knowledge about the collections of information with which they engage, and that is useful to record either for personal use or to be shared in a collaborative situation and to help in the correct interpretation of these information resources subsequently.

Knowledge Management systems have typically been presented as subsystems of enterprise applications likes Customer Relationship Management. The Nopik system is a standalone information system that supports the creation and organisation of information and knowledge inside a company and that can communicate with 3rd party applications to classify and retrieve information. To create and organize knowledge is a very general concept and which we have implemented through the use of three specific knowledge management methodologies for creating and organizing knowledge:

1. Support for remote access to process support environments: This supports mobile workers through web-based access to task lists through an e-work environment, where the user can download new jobs to perform and to upload the jobs completed.
2. Corporate Problem Solving: This is related to an e-management of non-conformity reports, where the user (employee or external expert) needs to report some anomaly or not conformity occurred on a plant or on an application managed by the system.

3. Web-based Corporate Information Management Environments: Supporting the creation of semantically structured and graphically navigable repositories of documents used by individual workers in the organisation and shared by workgroups and their annotation with knowledge records.

2. Objectives

In the emerging area of P2P computing, the concept of a "service" or an "interaction" is used in several ways. An important operational aspect of this relates to the perception of the final customer of that service. Only by the usage of such a service, is its operational relevance established; that is to say, the P2P network can be implemented either at the system architecture level, as has historically been the case, or at the user interface level which affords P2P functionality to the user without necessarily being implemented as P2P components.

The relevance of a P2P service at an operational level depends on the coherence of the constituent components. Breaking these barriers shall lead to true innovations in P2P service provisioning, but its success strongly depends on the basic architecture to be developed and its integral manageability. The aim of the work presented in the paper therefore is to derive a reference model for a dynamic networked enterprise consisting of communities (work groups), who share their de-centralized intelligence with other communities, thus creating a Virtual Community. Using plug-and-participate-technology, each one of the actors will be able to combine, share and manage their own resources and those found in community environment they are part of.

Finally, a key consideration is to keep centralized control to a minimum. The only “supervision” will arise from corporate Enterprise Resource Planning (ERP) systems that are concerned with the management of tangible assets and extend this to a new generation of systems that will enable the management of intangible assets to be included.

2.1 Remote Access to the Process Support Environment

The remote process support methodology requires the existence of processes, more or less structured, which users are required participate in and can take different forms:

- **Brainstorming**, a Name, Description, Keywords, List of People involved, where just a free Contribution form is required to be filled in. The first case could be also a consequence of Problem Solving or Innovation reports, where the system decides to open a discussion about a report, involving one or more experts.
- **Desktop**, where we can find a pool of diverse activities without any precedence relation, but each one associated to some precise application (i.e. tool + template) available for giving contributions (Office tools, Diagnostic Tools, etc.).
- **Project**, where each activity is part of a plan and has a start/end date, resources' planned/actual efforts, identifiable input/output documents, a precise application (optional) necessary to perform it.

In such a situation (see Figure 1), the user connected to the system has to carry out the required tasks. Through the Middleware layer or connecting directly using the web GUI the user retrieves his TO-DO list. The TO-DO list can be thought as a set of brainstorming, Desktops and Projects the user is required to give contributions to. In Brainstorming, the user is just required to fill in a web (or local) form with his contribution; in Desktop, he can use different tools (one or more) that are initialised to some standard template; in Project, he has to produce specific outputs as required, using inputs coming from previous activities. All the material necessary to produce the above GUI has to be downloaded from the system at the beginning of a working session; all the results produced have to be uploaded to system at the end of a working session.

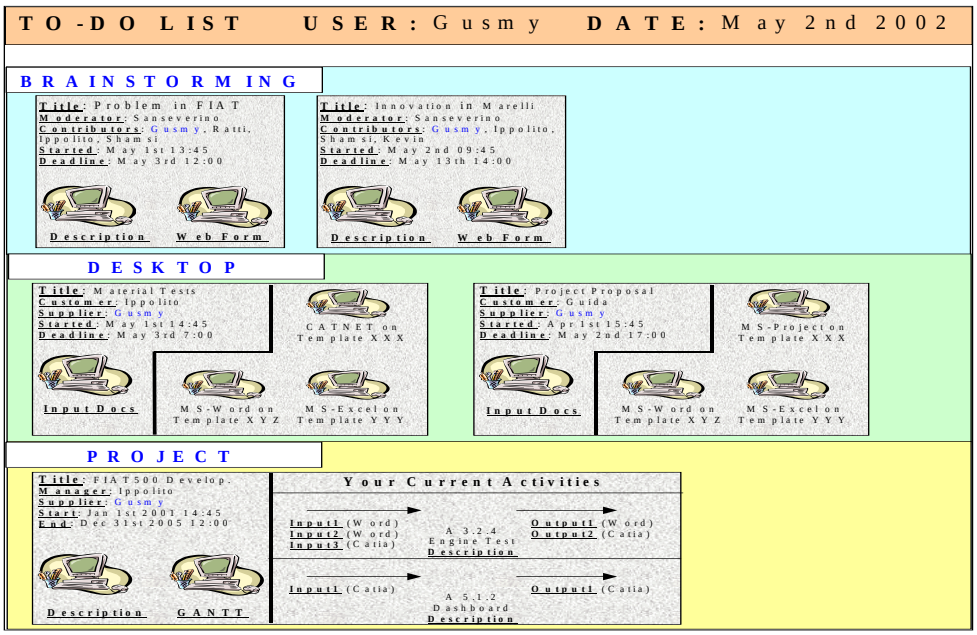

Figure 1: GUI for To-do List.

2.2 Problem Solving

In order to accomplish this task, the system module (accessed via a browser) is able to detect anomalies and to report an anomaly or not-conformity to one of the assigned servers. This module works independent of the pre-defined tele-working environment that organises user input through the brainstorming, desktop or project framework, however the user can access it and initiate problem analysis using a dedicated case based reasoning environment.

A new "Problem" can be input into the case based reasoning (CBR) tool using a web-form, where all the parameters needed are presented to the user. The CBR tool is set up for a particular application domain by specifying a domain specific ontology that in the case of non-conformity analysis contains details about products, customers, faults and alarms. The ontology also contains a number of different types of relations such as IS-A relations in the case of Customers, Faults or Alarms, or PART-OF as in Product. Each ontology description contains structured fields, unstructured fields and attachments.

3. The Use of Ontologies for Information Management

In terms of managing information and information retrieval in a flexible and semantically sensitive way there are important recent developments in the area of the Semantic Web that are relevant here as well as results from past and current framework projects. Thus work has been done on automated information extraction and classification techniques using ontologies as a classification principle currently developed in association with the SHOE and Parka projects [1] [2] [3] and others in the ontology community [4] including KAON, to name but a few. The basis of the work on integrated information and knowledge management paradigms developed as part of the Burma-X and NOPIK FP5 projects is reported in [5].

Ontologies are increasingly being used to underpin information systems but have in the main been used in the background to facilitate information storage retrieval and manipulation tasks or to provide thesauri for example in information retrieval systems.

One practical problem is how to arrive at useful ontologies for this purpose in the first place. Falling short of users generating them from scratch there have been a number of proposed approaches for the automatic generation of ontologies based on typical samples of

information from the domain in question such as typical documents. A considerable amount of work has been done on the techniques and tools are emerging such as [6]. A more comprehensive survey of available techniques was carried out in the course of the OntoWeb project and which highlights a considerable number of techniques and tools [7].

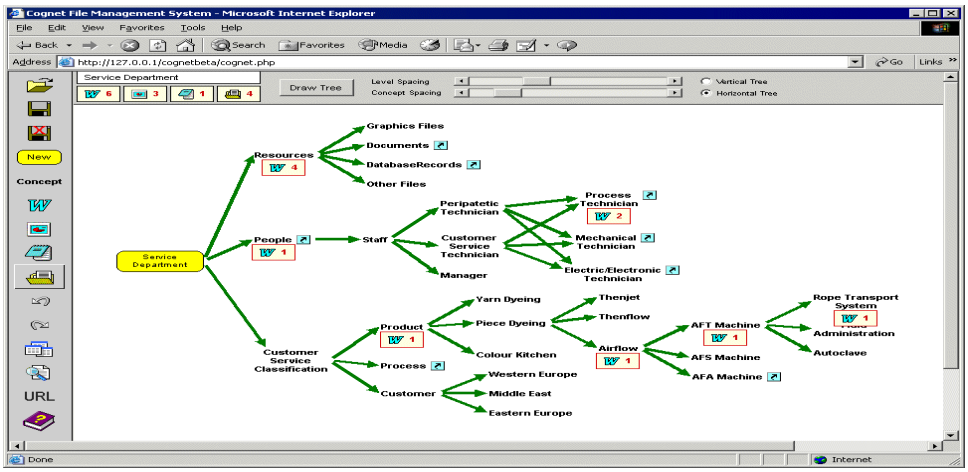

Figure 2: The Main Cognet User Interface

However, specific ontologies are usually constructed by analysts or ontology experts for a given use case and then used in an application and where the applications can vary in how they accommodate changes in the ontology as the application domain evolves over time.

Typically, the detail of the ontology is hidden from the end user or only parts of it are made visible. Furthermore users are usually not able to subsequently change these ontologies to suit their needs and the evolution of the domain or application in question. The work carried out in the NOPIK and Burma-X projects aims to reverse this trend and offer the user more control over these matters and to use them as the navigation principle. Thus the focus is on the use of ontologies at the user interface level, where users interact with the ontology not just to view results but evolve these ontologies to customize the tools to their domain of application. The evaluation of this approach to date has shown that the users are quite capable of devising meaningful ontologies and to use them to associate documents with them. Users have also responded positively to the use of graphical representation and navigation techniques that were chosen because they were expected to be more intuitive to the user than tabular representations. Figure 2 illustrates this and shows the main user interface with the graphical representation of the ontology used as a classification principle to which documents can be associated. Some of this innovation has therefore already being piloted in the Burma-X and NOPIK projects but work needs to be completed to provide a deeper use of ontologies and the more flexible use to manage a wider range of information types and additionally to record knowledge about the documents in question and their relations and make this fully navigable.

3.1 Ontology Impact

The proposed approach strongly relies on bringing ontologies to the user interface level and allowing users to engage with the ontologies and in particular with their visualisation in order to browse, store and retrieve content in a variety of formats and locations as well as knowledge records that exist about this content and to evolve the ontologies as the organisation or application domain changes.

As such, there is a strong element of using ontologies at a user level and on interactive visualisations. Amongst the currently available approaches and technologies there is still a lack of support of exactly this type of use. Plenty tools for ontology editing exist and which can be usefully re-employed and there are a variety instances where visualisations are

emerging such as OntoViz for Protégé (http://protege.stanford.edu) or hyperbolic graphs and 3D graphs now in use in commercial products such as OntoBroker and SemanticMiner from Ontoprise (www.ontoprise.de). Visualisations are also emerging elsewhere such as VisioDAML (www.daml.org) and Aidministrator, to name but two. However the visualisations are generally suffering from the problem that they are either to complex to be easily navigated by the average user or usually do not allow the user to interact with the visualisations other than to browse and retrieve; thus not supporting editing of them and using them to record and store. As for the support of ontologies at the user level there is still a big gap in the available technologies that continue to be aimed at the domain expert or expert user to build applications that use ontologies heavily as their backbone but not at the user interface level. In this respect the proposed environment that will continue work done in Burma-X (GRDI-2000-25560) and Pikon (renamed as NOPIK IST-2001-33487) and which will provide a much needed solution also for the semantic web if that is to focus substantially also on the human user level rather than the automated processing level.

3. Developments

Conceptually the NOPIK toolkit can be divided into a client and a server part, without necessarily implying a centralized architecture. The important aspect is that the users should have off line capabilities. However the existence of a server for the coordination and the synchronization of the content is considered useful and important.

The users access the environment and individual tools via a browser and the integrated environment offers a set of functionalities to the user for creating and organizing information and knowledge as well as problem solving related services and activities. The toolkit consists of seven components under an integrated user interface [8] including the following components:

1. Cognet (Ontology based information management environment)
2. Case Based Reasoning Module (CBR) for problem solving
3. A Search Module
4. Performance Indicator Module
5. Messaging Client
6. Problem Solving Process Manager
7. Office and OS automation components

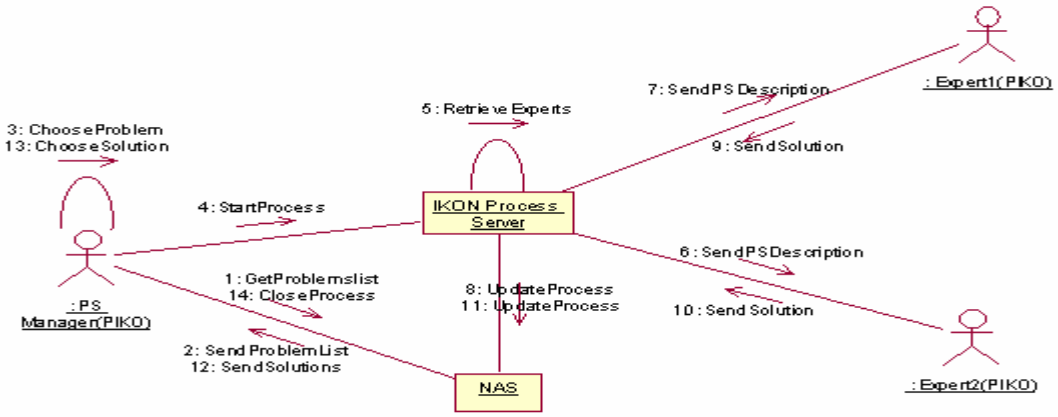

Figure3: Problem solving accomplishment for the single server case.

All of these seven components use a common repository where the information is gathered. This repository serves as database and document repository for the PIKO client (Personal Information and Knowledge Organiser). The client provides off-line support intended mainly for individual users to ensure that their personal resources are always

accessible even when working on the move. The IKON server (Information and Knowledge Organiser Network) consists of the Network Application Server (NAS) the Messaging server and the central repository of shared information resources. The NAS is the central module responsible for the synchronization between the different user environments and the individual servers. The Figure 3 shows in a collaboration diagram the communication between the users and one server instance in a Problem Solving scenario.

4. Adoption of the System

In a majority of enterprises or organisations considerable collections of information exist in document form but in many cases no systematic care is taken for managing information and knowledge as well as the flow of knowledge. Thus there are a considerable amount of under-utilised resources that can be potentially turned into commercial value if only the means existed to exploit them more proactively. However these collections of information are frequently under the control of individuals who have accumulated them often over a period of years. As these collections are growing they become unmanageable even for the user who accumulated them (let alone as a resource for the organisation). In the course of the NOPIK project an environment was developed that allows the user to use self-generated or imported light ontologies, that represent domain specific taxonomies to manage such collections and attach documents to them with the use of an automatic classifier and where associations can in addition be customised individually by the user. This has created nodes or islands of information and sense-making, but where the focus is largely on individuals or groups working with their shared taxonomies and shared collection of documents. This is designed for more explicit collaboration and which is supported by the tools developed in the course of the project. However, there are in addition frequently occurring situations where other users from within or without the organisation would want to gain access to this information in an ad hoc fashion and which is not easily supported in the current framework – a situation which will be typical of a semantic web world where the individual user may search for information he does not necessarily know (or wants to know) where it resides nor how it can be accessed.

The most appropriate architecture for open, ad hoc exchange of information is a peer-to-peer architecture where nodes that hold resources in a distributed fashion can connect to other nodes to allow access to their own resources and seek information from other nodes. Peer-to-peer networks have attracted a lot of criticism and mistrust because of copyright violations and a consequent reluctance of the commercial world to work with them. The question arising is therefore how we can harness them to ensure proper use and to ensure that privacy and access is properly managed. The key here lies in providing in principle universal connectivity while restricting access to and transfer of content based on, say, authentication and permissions. While in the case of Napster users were prepared to infringe copyright of music tracks in exchange for getting access to someone else's tracks, there is a difference with respect to personal and organisational collections of information in that there is on one hand a much more keen sense of ownership as individual users tend to be more protective of the information they consider key to their work and role in the organisation. In addition, there are ways and means of exerting control and auditability and which in fact would be greeted with more enthusiasm by prospective users that can thus exert control over the access and proliferation of their personal collections and to be able to audit and demonstrate the utilisation and hence utility of their resources and the added value they create. In this view the P2P platform will be the place where thematic features and spaces for the exchange of ideas and creative activity will be combined, and the needs of teleworkers and teams catered for. We refer to these islands as Basic Independent Knowledge nodes (BIKs). Each BIK is able to carry out the task for which it has been trained, within a given context of work, and without explicit in collaboration with other

BIKs. At any time a new BIK or some new functions of a BIK may be integrated into the system without any changes to any other BIK in a plug-and-participate fashion. Each BIK brings to the P2P network an added and innovative value, in terms of specialised services provided to the community.

There are several innovative things coming along with our system compared to existing solutions in the arena of Knowledge Management systems or approaches. First of all, our system is not only a toolset, but also a methodology. This methodology will ensure that the base for a successful KM environment within a company will be given namely by providing organisational methods in order to implement relevant organisational changes, which finally results in an environment where Knowledge is accepted as a value that has to be shared between people, whereas the intention here is not to force people but to convince them to share their knowledge with others.

The second innovative point is a technical one, namely the integration of an ontology based approach for the modelling and representation of knowledge, with a Case-Based Reasoning component in order to support the user while solving a specific problem. Although CBR could be seen as a method to represent knowledge on his own, the ontology-based approach coming along with graphical tools for modelling, representing and browsing through the knowledge base, is more users friendly and therefore especially suitable for domain experts.

However, for the day-by-day business, employees don't just need methods for the retrieval of knowledge, but also support for problem solving or decisions making. This should be based on the past experience of users in order to guarantee best results. However, this more "process-based" knowledge will be kept within the CBR component and therefore, both approaches will complement each other.

In addition to the things mentioned above, there are some other features of the system that make it different from existing tools. The web-based architecture will ensure that the tool is applicable in arbitrary IT environments, and furthermore reflects the distributed enterprise of today. The resulting system is composed of different components, which addresses issues regarding scalability and the cost to developer and eventual user and that is of key importance especially for SMEs. Finally, the solution also addresses the needs of domain experts, which means that every user (a blue- or a white-collar worker) will be able to use this tool in a user-friendly way in order to share his knowledge and to improve the competitiveness of his company.

5. Conclusions

The material presented is based on tools that are being developed in the course of the NOPIK project and represents a proposition for how this could be extended in future through a true P2P architecture and the benefits that can be expected. The experience with the basic problem solving architecture in the course of the project has been positive and its capabilities to substantially support problem solving activities has been demonstrated in user trials. A number of lessons have been learnt, including that the analysis need for the introduction of customisable solutions needs careful attention not because the customisation per se would be problematic but because it requires active input from the eventual user who may not be accustomed to analysing the way business processes are carried out or to changing them readily even though in principle it would be beneficial to the organisation due to organisational constraints. However though the tools need customisation test users have largely been able to develop categorisations for the Cognet tool, for instance, with relative ease but unfortunately at the time of writing, the ontology learning techniques available can't be applied automatically.

The work of Cognet has been to evaluate the ability of ordinary users of engaging with ontologies and has as such been a success. The technology including the CBR tool and

several other tools besides has been tested with three main users in the course of the project and with a number of additional voluntary users. The use cases include managing business consultancy services of a regional development agency, a customer complaints handling process from a construction sector supplier and anomaly handling in the automotive sector.

Problem solving is a time and resource consuming activity in organisations as for example in complaints handling and customer service and where concrete economies can be achieved by utilising suitable information management and decision support techniques as well as streamlining the involvement of experts. The P2P concept presented is a proposition of how to extend the functionalities of the NOPIK toolset and to further exploit the knowledge management and open interchange and sharing paradigms that allow more ready connections between distributed nodes in an environment where participants may need to connect in an ad hoc fashion and where current typical technical incompatibilities readily limit interoperability. Consequently the material presented in this paper, while based in part on existing work, will be the basis for future work in order to generate the P2P platform that will need to be superimposed on the existing technology. Early work with this has already happened through the distributed architecture and networked application server developed, but which will need to be extended to generate the full functionality that has been argued for here. Proposals for extension of the work as a whole and commercial exploitation avenues of individual tools are currently under investigation. At present an advanced demonstrator toolset is available as the NOPIK project is in its final stage and which is web based and includes the synchronisation of distributed installations but on a static basis rather than in a dynamic, ad hoc fashion as would be the case in a true P2P scenario.

The proposed approach aims to contribute to the research and innovation in the field of open integrated management of information and knowledge using ontologies and semantic web technologies that and P2P computing. More specifically, the proposition is to introduce a novel P2P platform for facilitating rapid service creation and maintenance, so that networked enterprise processes can be established on an ad hoc manner, i.e. dynamically and on demand.

Acknowledgements

This work has been partly funded by the European Commission IST Programme. The authors wish to gratefully acknowledge their support as well as the contribution of the NOPIK project partners.

References

[1] J. Heflin, J. Hendler, In: “Dynamic Ontologies on the Web”, Proc of the 7th Int. Conf on Artificial Intelligence (AAAI-2000), Menlo Park CA/Cambridge MA, AAAI/MIT Press

[2] J. Heflin, J. Hendler, In: “Searching the Web with SHOE”, in Artificial Intelligence for Web Search. Papers from the AAAI Workshop, WS-00-01, Menlo Park CA, AAAI Press

[3] J. Heflin, J. Hendler, In: “Semantic Interoperability on the Web”, Proc of Extreme Markup Languages 2000, Graphic Communications Assoc. AlexandriaVA

[4] S. Staab, J. Angel, S. Decker, M. Erdmann, A. Hotho, A. Maedche, HP. Schnurr, R. Studer, Y. Sure, In: “Semantic Community Web Portals”, 2000, Computer Networks 33(1-6) pp. 473-491

[5] A. Slade, A. Bokma, In: “Ontologies within Extended Enterprises”, Proc of the 35th Hawaii Int. Conf on Systems Science, Hawaii Big Island, January 2001, Computer Society Press

[6] Maedche, A. and Staab, S.: The Text-To-Onto Ontology Learning Environment. Software Demonstration at ICCS-2000 - Eight International Conference on Conceptual Structures. August, 14-18, 2000, Darmstadt, Germany.

[7] A. Gomez-Perez and D. Manzano-Macho, “Deliverable 1.5: A survey of ontology learning methods and techniques”

[8] D. Tektonidis, R. Ratti, N. Solomon, A. Bokma, L. Rabe, In: “NOPIK Generic Integrated System: Software Concept, Architecture”, Deliverable D3.6, Thessaloniki, Greece