CognIT a.s

Rapport

Creating a Learning Organisation Through Content Based Document Management

By
Bernt A. Bremdal
Fred Johansen
Christophe Spaggiari
Robert Engels
Richard Jones

CognIT a.s

Introduction

The discussion on the concept of the Learning Organisation dates back approximately 20 years in the management literature. People that pioneered this concept include Chris Argyris (Argyris 77), Peter Senge (Senge 90), Fiol and Lyles (Fiol 85), Levitt and March (Levitt 89), Ray Stata (Stata 89). All of them introduced various definitions of the concept. These definitions circumscribe issues like the following: "Better knowledge and understanding", "process of improving actions", "processing of information", "change of behaviour", "encoding inferences from history into routines that guide behaviour", "process of detecting and correcting error", "the need for shared insights, knowledge and mental models", "building on past knowledge and experience". Peter Senge's work (Senge 90) is often used as reference for the concept. His ideas put forward in the book The Fifth Discipline have had a profound effect on modern organisational thinking far beyond the management community itself. According to Senge a Learning Organisation can be described as follows:

"A Learning Organisation is a place where people continually expand their capacity to create the results they truly desire, where new and expansive patterns of thinking are nurtured, where collective aspiration is set free, and where people are continually learning how to learn together".

Organisational learning must be seen as part of an overall process of continuous improvement. But practical and systematic measures must accompany a vision if a real change is desired. Garvin (Garvin 93) is concerned about the lack of an operational basis or framework that can systematically instantiate the concept into a viable strategy and a set of systematic actions. He identifies five properties that learning organisations master well: Systematic problem solving, experimentation with new approaches, learning from their own experience and past history, learning from the experiences and best practices of others, and transferring knowledge quickly and efficiently throughout the organisation. Although most organisations handle all of these to some degree few approach all of them in a systematic manner. Garvin attempts to define a platform that can support the definitions of systematic and practical actions. Recently we have seen the introduction of new technologies such as the Internet and the Intranet that may provide parts of such a platform. However, this may not be enough. In the TEKMO project we have explored a concept that capitalises on Senge's vision and Garvin's ideas.

A knowledge-creating organisation relies on intellectual strength and information in order to achieve its objectives. Previously we have referred to such companies as "Information Refineries" (Stokke 94). Examples of such are hospitals, engineering companies, high-tech companies and financial institutions. Electronic document processing has become the order of the day for most of these. Even historic material written on paper can be conveniently transferred to electronic means. Information refineries produce documents continually as part of their main stream work in order to support long-term planning, provide advice, make decisions or conclude a production process in the form of drawings and specifications. In the present discourse we treat the document as the bridge between established work practices and a new form of communication that needs to drive the Learning Organisation.

According to Feldman (Feldman 98) 80% of all available information appear in non-structured collections of text drawn from various document sources. Despite its ill-structuredness and other obvious weaknesses we claim that documents represent the organisation's best knowledge repository beyond the people themselves. Since companies have established routines for document production and processing of diverse nature we suggest that any knowledge management effort try to capitalise on this. This implies the use of new publication techniques provided for by the HTML-based Intranet. It also calls for ways to structure and systemise the content embedded in the different documents in order to elicit knowledge for both problem solving and learning across the whole organisation. We will refer to this as content based document techniques. Content based document management enables knowledge and information accommodated in all documents, despite quality and size, to be systematically indexed, retrieved and fused. Given an electronic network it is thus possible to create a liberal flow of knowledge across one end of an organisation, despite departmental boundaries, in the form of an indexed document or explicit descriptions of the document semantics in the form of ontologies. Work efficiency can be maintained while people are able to interact and exchange knowledge since both efforts are focused at the document. If, in addition, it is possible to organise interaction on the net so that people really exchange ideas, documents and knowledge the possibilities of a Learning Organisation will become real. For this we look at established organisational structures and try to blend this with an electronic infrastructure.

Organisational problem solving and learning

Using Garvin as a guide we may claim that learning organisations are skilled at creating knowledge, acquiring knowledge and transferring knowledge. This requires a high degree of awareness and perception. Knowledge advances are propelled by new ideas encompassed by the data flux surrounding the individual and his organisation. However, knowledge alone is not sufficient for change. An organisation that learns must readily respond according to its knowledge. This is essential for all types of intelligent behaviour. On his discussion on bounded rationality Herb Simon (Simon 76) points to how limited knowledge governs our behaviour. The familiar story of the berry picker may serve as an illustration of this influence.

A person's goal is to pick as much berries in the woods as possible within a time frame. He is made aware of the fact that there are a significant amount of berries at a particular spot in the woods. He goes there and fills his buckets and return without knowing that behind the hill, close to the spot where he went is another area with far more prime quality berries. As watchers we cannot blame him. With his limited information he acted in a rationale way. Let us assume that he, in fact, was aware of the better spot. This place is sufficiently accessible to pose no extra demands on his available resources. Provided that he maintains his original goal we would probably judge his act as irrational. The morale is simple and applies to an organisation as well as an individual. If the knowledge is available exploit it and maximise the effect of your actions.

Problem solving and learning often appears as distinct and only remotely connected. Aamodt's work (Aamodt 91) may serve as a good, model-based round up of the contrary. Learning is often a product of problem solving. Problem solving requires continuos learning. For smart individuals this work well. In spite of this an organisation may have significant problems of bringing the two together. In fact, the tight connection between these two principal aspects of cognitive behaviour serves as a grand obstacle to the organisational aim. What is learnt in one end of the organisation is not readily available at another. Consequentially we may find that problem solving undertaken in one project may not apply experience gathered in another one. Things tend to initiate from scratch and painful issues tend to be recurrent because of this. Reasons for this stem from the fact that people operate within a work environment that does not encourage what we will refer to later as "real communication".

Flexibility and adaptability are essential virtues of a learning organisation. It must modify its behaviour to reflect new knowledge and insight. Unlike the individual, however, an organisation is not borne as a single bodied entity. This implies that the awareness aspect, the knowledge creating process and the action part must be brought together by external means. The advent of new computer technology and the introduction of the Internet/Intranet promise a change to this. Even extremely large and distributed organisations can be linked together. However, this alone does not solve how the different cognitive processes distributed among individuals at different ends of the organisation can be unified. Historically we have seen the impact and the limitations of the telegraph, the telephone and other media. The real issue is how to build effective cognitive processes on top of the network. In a sense we are looking for an architecture and a control paradigm that in an organisational perspective resemble the human nervous system. A slight touch of the finger on a sharp edge will cause an immediate reflex in the central nervous system, take control of all resources and make other limbs step back in a unified manner.

Management and work practices

The CommonKADS Model

Several theses have been published on the need for flatter organisations, cross-functional teams and improved interfacing between professionals in adhocratic organisational structures (Mintzberg 83) Yet, when it comes to day-to-day practice efforts to reach such goals are often hampered by traditional management values and attitudes towards work. Efficiency should not be sacrificed for improved interaction between people. This issue can be described through a common value chain perspective using the Common KADS model (de Hoog 94, Schreiber 98) (see excerpt in figure 1). We will explain the importance of this model. The value chain perspective applies a process-oriented view of the organisation focusing on how input acquired by the organisation is gradually turned into a result representing a higher value than the input. A process will consist of a series of tasks that each contributes to the increase in

value. An agent is responsible for one or more tasks. An agent is often a person. It might also be a machine. More and more we see that an agent is the combination of both a human and a machine. Nevertheless the main objective is performance. Maximise output with the least amount of resources. Hence an agent will be tuned or trained for maximum accomplishment. Things that take the attention away from the performance aspect are viewed as counter-efficient. Historically the communication act was limited to a set of simple control signals from above, "start-work" and "stop work". Orders of this kind characterise the military hierarchy and used to be the standard for communication within most organisations until the last half of this century. Gradually more communication has forced its way through, but it is still seen as mostly an ad hoc activity that must be minimised. The communication task that the learning organisation must pursue should include transfer of knowledge. This is what a substantial part of the CommonKADS effort has been about (Schreiber 98). Government and business professionals are positioned to tackle ever-growing complexity and dynamics. However, the problems that they are faced with can no longer be solved on the basis of one person's skill and know-how alone. This calls for organisational structures that encourage effective knowledge sharing and knowledge building.

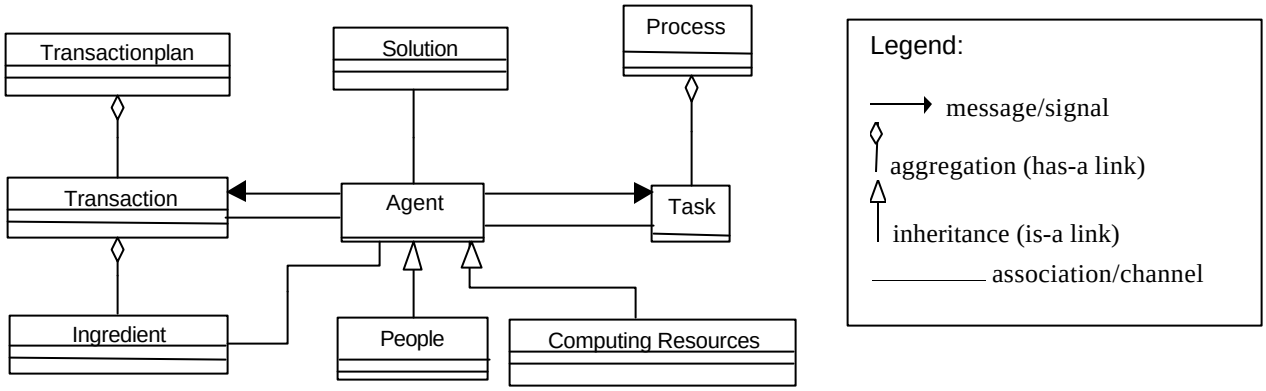


Figure 1. The combined responsibility of performance and communications as depicted in the CommonKADS model

Efficiency versus communication

Davenport and Prusak (Davenport 98) have pointed out the inherent knowledge exchange effect caused by informal conversations such as the "coffee machine talk". They also dwell with the low regard many managers hold for such informal exchanges. Despite their knowledge transfer capacity and potential for cross-fertilisation of new ideas informal communication has not found its place in the efficiency focused and performance oriented company. Another example is the staff meeting. The slogan "no partying, decisions only" has been proposed in good will in many professional organisations in order to cut time spent in meetings. Yet Bremdal (Bremdal 98) claims that the performance objective is killing the educational aspect meetings. People are not allowed to talk freely and discuss company or project issues within the framework of a staff meeting or project meeting. People that seldom ever meet and that have different backgrounds need time to socialise on a professional level and find a common language. Within the stringent framework of the westerly type of meeting culture there is no place for a gradual consensus type of movement. Quick decisions and back to work have been hammered into us more than ever. A slight departure from this is seen in problem solving type meetings often found in engineering companies. Sometimes this is the prime knowledge exchange forum in software engineering projects. This is where the functional responsible people explain to the programming part of the team the concepts and the rational of the business processes that the new software should support. This is where the computer people pass on their ideas on how functional aspects can be handled by novel computer technology. However, we have also observed how the project team gets stretched under the pressure of time limits and other concurrent demands on individuals from other projects. People fail to show up to meetings and the educational aspect of the assembly falters. Sometimes the meeting is totally relinquished on the alter of seemingly concrete action. The emergence of e-mail is beginning to change a little of this due to its asynchronous nature of communication. E-mail has become a new medium that supports both formal and

informal exchanges. In a sense a project meeting and the "coffee machine talk" can be transferred to the net.

Corner stones that support the communication act in Learning Organisations

When defining a practical communication plan for the Learning Organisation we can look back at historical initiatives. Formal approaches have been introduced to enhance and manage this. Two corner stones that we have brought to bear are "communication transparency" and the adhocratic organisation.

The main ingredients in the communication plan must always be decision oriented. But the result of the decision process is not sufficient to support neither learning nor problem solving in groups. It is important that the decision process is made as transparent as possible. Rittel advocated design through argumentation almost thirty years ago (Rittel 72). His main point being that a complex decision process like ship or car design must muster a broad basis of expertise and information. This basis is too large for any individual to master alone. Only through extensive exchange of ideas and the arguments that support these are we able to grasp the implications and solicit the alternative that will determine the actions of the company. He advocates sincere involvement and to make every voice heard. Arguments make things explicit and issues that are explained create a shared understanding of what can work and what will not. A more formal approach to the debate issue has been pioneered by Mikkelsen et al (Mikkelsen 96). This initiative shows that computers can capture the ideas of Rittel within a formal, forsenic type of framework.

Mintzberg addresses the adhocracy (Mintzberg 83). This is the organisational structure that handles problem solving with groups of people representing different professions. An adhocracy is like a project organisation poised to handle problems. The only difference being that the organisational structure dominates the whole institution. Several successful adhocracies have proven its worth. NASA, during the sixties and early seventies, may be the largest and most successful adhocracy in history according to Mintzberg. Extremely apt at solving problems and at re-organising to face dynamics the pure adhocratic organisation is not very efficient at handling ordinary, routine things. The root of its inefficiency is the adhocratic organisations' high cost of communication. "People talk a lot in these organisational structures, that is how they combine their knowledge to develop new ideas".

Clearly there must be a balance between performance and communication. We may claim with some historic support that companies have spent more time becoming more productive through various efficiency measures than looking at the communication act. Machines have been introduced to support production. But has the communication act received as much attention? Issues related to too much talking, being around the coffee machine or in the project meeting, may well be due to the fact that we do not know as much about real communication as we should like to. Few people are really trained in making their ideas transparent through proper argumentation techniques and publication. The inefficiency of a giant adhocracy like NASA during the days of the Apollo and the Gemini may not be a problem of organisational design, but lack of optimisation. What if NASA had e-mail at the time? The initiatives behind the Learning Organisation represent the first steps towards an organisational structure that does not sacrifice one for the other. Efficiency is still a goal, but efficiency on the account of the whole team, not only the individual. The communication process must support this. The work that we present here acknowledges the fact that the latter requires technical support. Our focus has been to develop a system that can create a balance. The means for achieving this is the product of the understanding of what "real communication" is and to seek solutions that build on established document practices. The reasons for this are obvious. Document production is a key task in the knowledge-oriented company. The communication act can readily be built on top of the production task. Written material, formal or informal represents a persistent externalisation of some knowledge. In addition it is more structured than simple talking. Network technologies can bring people together across great distances. What we seek is to maintain the efficiency of a well-trimmed production company with the problem solving and learning capabilities of the adhocracy.

Real communication and shared ontologies

In the TEKMO project we have been concerned with real communication. By "real communication" we mean to what degree a receiver is able to capture the essence of the sender's mental concepts. This demotes the value of signal volume. It strongly emphasises the ability of the sender to express his mental model in a form that stirs a resonance with the receiver enabling him to expand his mental model,

reinforce it or reject the input. A good match between expression and concept increases transparency and enhances both argumentation and the learning effect. Our approach can be explained in terms of Shannon's information theory (Shannon 48) and his discussion on information redundancy. A set of carrier terms occupies a central position in this philosophy. Carrier terms are tokens that define a context and even the goals of the communicators. In order to establish real communication the speaking or writing part need to overcome two principal challenges and to combine the two:

- To identify the terms that express and convey, within a given context, the mental concepts in his own mind.
- To judge what tokens would best match the conceptual models of the audience.

Communication is clearly a very knowledge oriented effort and has been referred to in terms of ontologies in different types of literature (Genesereth 92, Gruber 93, Hamon 96, Swartout 99). Ontologies can be defined as a structure of tokens that portrays a mental model.

The TEKMO project has focused towards shared ontologies. Those are ontologies that are communal and which we claim to be a prerequisite for real communication. The basic philosophy driving the project has been that documents embrace and hide ontologies. Although incomplete and sometimes erroneous these ontologies contain the carrier terms that determine the link between the author and the reader. Our theory is that a well-written document although unstructured in form may express terms in a manner that immediately triggers the interest and understanding of the receiver. Observations from tests with human subjects support the claim that given a context a set of tokens can be sufficient in order to establish a mutual understanding between two people. Ontologies describe contexts. The ontological aspects of a document are essential for stimulating the awareness of the receiver and his degree of perception. This is based on what Shannon termed linguistic redundancy and entropy. Given a term we may anticipate a set of other terms in a particular combination. Moreover the structure of an ontology may articulate even tacit concepts. People are able "to read between the lines". If the context is apparent from the situation or the text a reader will actually add more than is precisely written. The contextual awareness and the ability to combine concepts that are actually articulated in the text drives this interpretation process. Our claim is that ontologies extracted from text can even express this type of tacitness. The work in TEKMO has thus been to investigate how we can build ontologies that capture the essence of a domain and use this in order to support communication. Indeed our focus has been to build a system on top of standard document routines in order to amplify and enhance the role of documents to meet the requirements of the Learning Organisation.

The importance of the Intranet and Internet

Collaborative efforts on the net

Over the past few years we have seen an explosive interest in Internet that is closely followed by corporate interest in Intranet technologies. The Intranet serves the corporations like the Internet does for the world. It may link people together in ways not seen before and reduce the importance and need of a company's middle management (Drucker 88). Modern computer technology can improve accessibility of people and their work, magnitudes better than the telephone has managed.

Reports from Shell (Speh 98) implicates that their Intranet constitute an important medium for accessing expertise and seek advice and to find relevant best practices and knowledge pertinent to a situation. At Hewlett-Packard they see the Intranet as an important means to link together a decentralised organisation (Gannon 98). The Intranet has become an instrument for individuals to seek support and work better together. Hewlett-Packard's knowledge management strategy aims towards a use of Intranet that "reflect the culture from its entrepreneur days". It allows the people at HP to initiate and grow changes from the bottom despite its size. Like Garvin, Speh argues that organisational learning has been subject to too much abstraction and conceptualisation without the support of relevant technology and without sufficient emphasis on structured knowledge. To him and many of his peers the Intranet constitute the right infrastructure for a Learning Organisation. The Intranet may have a profound effect on how a social organisation works and behaves.

Interaction is not limited to advice or occasional calls. Very goal oriented collaborative work efforts on the net have been reported over the years. Several initiatives both within and beyond the computer community have explored the concept and proven fruitful. One of the most extensive ones is the work on the Apache HTTP Server, a package of software for the public Internet market (Fielding 97). The project has been managed by the Apache Group and set out to build a piece of software. The Apache group was from the start a very geographically scattered group of volunteers with expertise in various aspects of the Internet. The highly successful product development was not managed by traditional means. The effort had no centralised centre of control or any formal organisation.

In search of proper stimuli for net interaction

All efforts demonstrating successful interactions on the net seem to be driven by intuitive ideas and situational needs only. The reports that we have on these serve as mere empiricism. However, Wilson (Wilson 98) claims that collaborative efforts on the net are closely dependent on the degree of interactiveness already established within the basic organisation. If the communication act is not well developed within the organisation we might expect the same isolated and non-social behaviour on the net. In organisations that maintain an inert style and where interaction between people is not so high the installation of an Intranet can be a disappointment. According to Wilson Intranet technology alone may fail to increase knowledge sharing. Desire rather than capacity drive social interaction alone. The necessary stimuli for proper interaction may still be lacking despite a high degree of connectivity. In contrast a very active organisation may add too much to the net swamping individuals in e-mails, bulletins, newsgroup updates etc. Control is one measure that can be added to reduce the overload. However, strict control works counter to the idea of maintaining open channels and a positive entropy. Another approach would be to aid the individual in navigating in a steadily growing knowledge plantation.

Due to the static of the net it requires active participation and a high degree of interactivity on behalf of the individual user in order to yield value. Thus the individual must be attracted to the net by recognised mechanisms that rule social behaviour and participation. In the case of the Apache HTTP Server the authors (Fielding 97) point to the importance of being well acquainted with e-mails and the net. The initiators shared the same profession. Fielding and Kaiser also highlight the need for a common goal and a shared information space for access of both current and past communication and development of artefacts.

Intuitively we may question to what degree people are willing to interact and socialise on the net. Adults, even professional computer people, still initiate a major part of their social and working relations in the physical world. However, an increased number maintains these relationships by means of electronic communication. For younger people Internet initiated relationships seem more frequent. New acquaintances and friends are made through chatting, e-mail and the use of home pages. Here is where people get acquainted with other people beyond their physical reach. People that share one of their hobbies, that have visited the same place, that are struck by the same serious illness as themselves, that have applied for the same school and that are working on the same class assignment as their class and so forth. The Intranet and the Internet may bring people together, given the right conditions for this. Until this point our basic hypothesis is that the basic stimuli for this is driven by personal interest.

The technical capabilities of the net are necessary, but not sufficient. But carefully exploited the potential for socialisation and exchange of know-how, know-why and other kinds of knowledge and experience for both problem solving and learning is there. Knowledge management must bring to bear incitements for network use and a language of communication that includes the people that need to be included. It must encourage increased interactivity to the passive organisation through the Intranet in order to achieve Senge's learning ambition. To the very active organisation it must add guidance and intelligent filters. In this manner the knowledge repository can render the effect that its potential promises.

The TEKMO project

Objectives

The aim of the TEKMO project has been to develop a system that can support the communication act that we have discussed here by building on existing document creation and document management practices. The Intranet constitutes the platform. It maintains the channels and the capacity for propagating stimuli like in the biological nervous system. A centrally controlled memory empowers the network classifying the stimuli according to previous experiences. Information flows back and forth in the system. Such transactions encompass documents of diverse variety. The document management aspect has circled around how to propagate the essence of the document rather than the document itself. The essence is defined as the product of document content, reader's interest and context. Combined with ways to encourage constructive interaction on the net and intelligently manage the social aspects our ultimate goal has been to empower the organisation with a systematic approach to learning. Figure 2 illustrates the concept.

Figure 2. Organisational Learning built on top of content based document management and Intranet technology

The system developed

In figure 3 we have depicted our system. The central unit is the information and knowledge server. Despite playing a central role in our system the operation of the Intranet does not depend on this. It is meant to play an adjunct role, more like a librarian than a manager does. Yet it must monitor the main stream of interaction in the network. The server services the organisation in three major ways. Those are:

- professional socialisation on the net
- publication support
- knowledge sharing, knowledge building and knowledge preservation

All are incorporated in the CORPORUM concept, but the TEKMO project is primarily concerned with the third issue. We will address the two first briefly. The third will be dealt with in more detail.

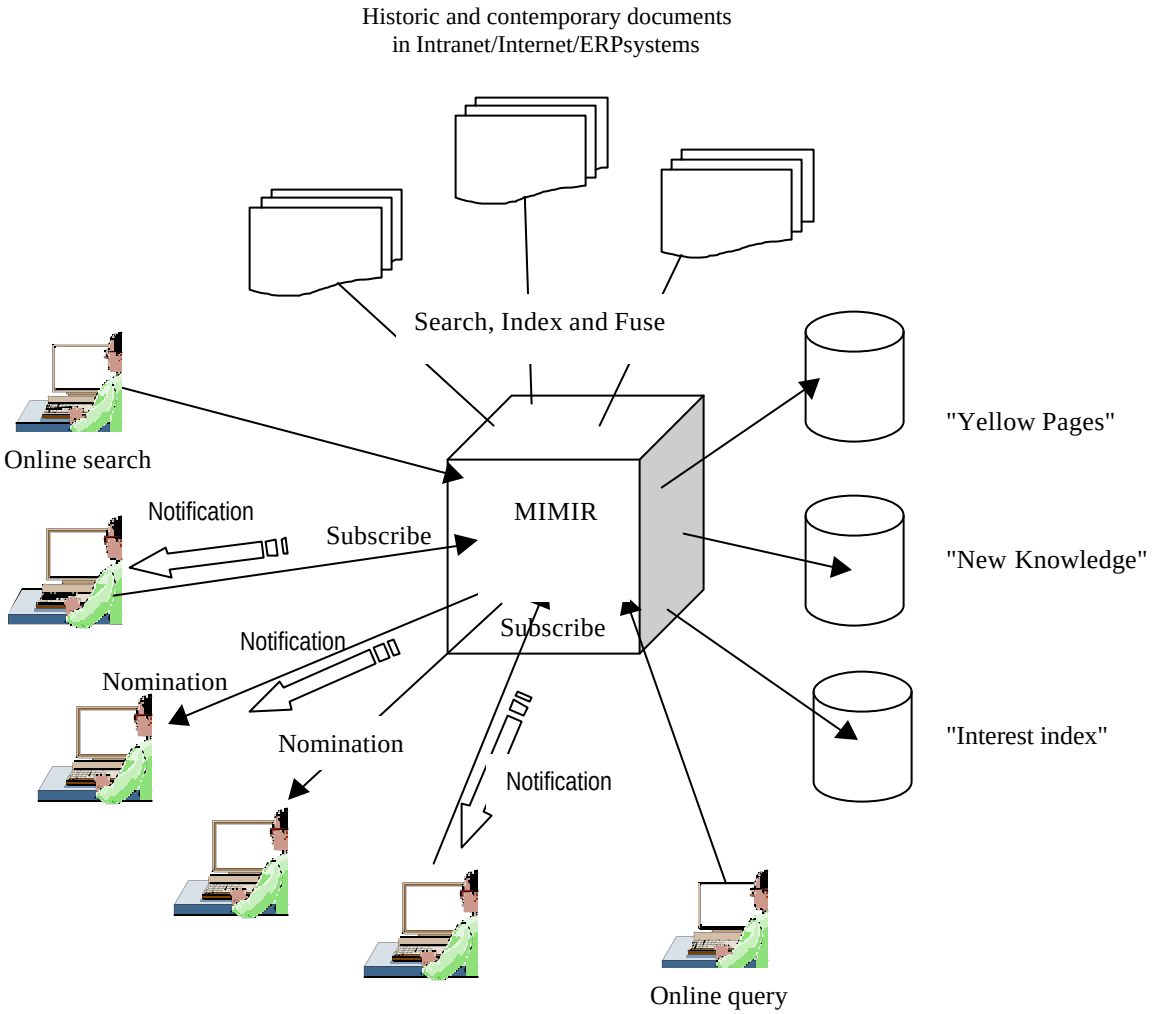
Software agents

An essential feature of the whole system is that it is agent based. On a technical level this implies that we apply computer programs that are encapsulated in separate and almost independent computer processes. Each program represents a particular user or user group and are both customised and created as such. A software agent has its own memory and controls much of its own behaviour. It responds to the environment according to predefined instructions or complex inferences stemming from such. On a conceptual level our agents represent the processing elements of a nervous system. The agents that we have defined in the TEKMO project are loaded with information about the net. Each agent has the capacity to enter areas in the net that requires special operations or access rights. Given the rights and the correct instructions by the user agents may penetrate deep into areas that would otherwise require multiple operations on account of the user himself. All agents in the TEKMO project are assigned an area of operation. This area may be a single Web node or a database. In principle it may be as large as the full Internet. Agents may also be given life cycle instructions. They may respond to events on the net, time and availability of data.

Using a set of instructions or rules they may decide to rest, work, communicate and even kill themselves without any interference from the user who created them. Agents may operate according to simple rules that define the interaction with their creators, other users and other agents. These rules give the agents certain social qualities. The agents created in the TEKMO project are also equipped with capabilities to interpret text, extract knowledge from this, communicate it to others and fuse output from one node with several others. An important property of these agents is the interest profile they contain. Complete or incomplete, they use this information to solicit material that is pertinent to a project, to a group or to an individual. The interest profile may take on three facets, that of the organisational role (i.e. group leader, project leader, QA-responsible), the interest associated with the profession (i.e. economist, systems engineer) and interests related to current activities that include problem solving tasks as well as learning.

The role of ontologies

The system developed in the TEKMO project is able to exploit both historic and contemporary material by extracting ontologies from documents and use this to capture relevant information, interpret it and pass it



on to addresses that can benefit from the knowledge that this represent. Articulated interests can be converted to ontologies and posted for subscription and search purposes. Those ontologies represent the receiver's mental model while the documents found and processed represent that of the publisher. Real communication can thus be established between people who may not be aware of each other's focus and concern. Only meaningful and contextual relevant data is passed to the receiver despite the increasing level of publication both on the Internet as well as in the Intranet. Given that a document reports the result of a working task and includes, at least partly, the rationale behind the conclusion our approach will piggyback the traditional task performance oriented business philosophy. But at the same time it will ensure a communication effort that relays context specific knowledge to problem solvers or learners that operate within the network environment. In that manner advances toward an efficient learning organisation can be achieved.

Socialisation support

As pointed out earlier adults, even frequent users of the net, seldom initialise social contact through their browser or e-mail system. But they tend to maintain such relationships by electronic means more and more. Younger people have come much further. Long term relationships, even marriages, are increasingly the product of Web surfing and chatting. Evading the sociological issues associated with this, we are content with the observation that both professional and emotional relationships can be established and maintained through the net alone. One prerequisite for this is that people feel invited and sometimes ushered into social interaction. Combined with a stimulus targeted at their professional and personal interests and guidance chances are good that people will grab the opportunity to interact. In the TEKMO project agents can take on the role of interest brokers and network ushers. Footprints of people's interest are left in all areas of the electronic world. Documents that are written accommodate this. What people read leaves similar impressions. Agents can detect this through their ability to interpret text and compare that with the interest models of other people. They may even detect other agents and query their interest, assuming that the interest model contained with them represents the interest of their master. Within a security framework that observes the integrity of both people and organisations these agents can establish shared interests and then initiate actions accordingly. Such actions can be invitations to join news group, problem solving teams and support groups for projects and sales teams.

Publication support

The CORPORUM system that we are building in the TEKMO project is designed to handle documents that are made available on the net regardless of its original purpose. The philosophy is that all documentation, even the electronic version of the "post it" labels should be handled. This is due to the fact that all historic material, even paper based, constitute a tremendous reservoir of knowledge. Historic material related to the core competence area of a company may represent a stable and potent source of knowledge that may be a sleeping capital that now can be put to work again. Nevertheless the Learning Organisation should strive to improve both writing and publication. With the advent of the word processor every person can now be his own secretary. Standards can be introduced to improve both the explicitness of theme, context and the availability of new documents. Rittel's argumentation concept can be manifested in a written format. We propose that companies apply practices from the conference world. Above all, authors must learn to address a wider audience. The future reader may well prove to be just as important to the company as the contemporary is. Document templates and wizards must be applied in order to support both outlining, summarisation and story telling practices. The latter is important in order to place a topic within a historical perspective, generalise language and make issues more explicit. Agents created in the TEKMO project can assist in linking the authors' work to historic material. Thus establishing a context for improved overview and contextual understanding.

The technology developed supports content tagging using the Extended Mark-Up Language (XML). Agents will help the user to enter HTML like tags that define the content of a document. This classification is supported by a company-wide indexing system. The tags represent a data model that can be queried in a structured manner that enables both people and computer programs such as a Browser to gain access to data and knowledge including author, subject, prices, equipment descriptions, explanations, financial statements and contracts.

An Intranet often consists of a static part and a dynamic part. The dynamic part will be the most extensive and less accessible given the dynamics and the distributed publication responsibility. CORPORUM is able to create semantically based hyper links between the dynamic part and the static part. A long-term implication of this is that the traditional hierarchical archive will disappear. People will "click and skip" through the information world, not thumbnail through a computer emulated Kardex drawer.

Grabbing and searching for pertinent information

As pointed out above written specifications of interests can be interpreted like other documents. An ontological representation is thus established that reflects some of the issuer's mental concepts. This can be stored and maintained in different ways. The most common is to include it as part of the specification that an electronic agent needs to operate. Another would be to leave it as part of a user profile on a document server i.e.Web server. Electronic agents use this ontology to solicit other documents and see if there is a match. Partial matches are accepted. Placed on a server the interest model assists as a subscription mechanism. By monitoring the data flux interesting documents that pass by can be captured and relayed to the subscriber.

Subscription and notification

An inherent consequence of agent persistence and its ability to support the user with pertinent news and updates is subscription. Both periodic an event based updates are possible. With sufficient individual precision news updates like this can be extremely valuable. Results documented in one end of the organisation may at once be captured and relayed to a person whose work and actions are immediately influenced by this. The publisher will never have to manage the communication act alone. An individual can also increase his awareness through subscription to URL's on the WWW. Pertinent updates from professional news agencies such as CNN, incremental changes to company news pages, new additions made to a list of published patents and updates made on behalf of universities and government agencies are a few examples. Subscription focusing at an individual's particular interest, combined with immediate notification enables people to engage in a day to day learning activity of great significance without limiting his basic working performance. Notifications can be presented in terms of e-mail sent from the agent to his master or as automatic updates on the user's private "readers list".

Indexing documents

Using the key concepts in the ontology documents can be indexed according to content rather than using ad hoc words pertaining to a foreign archive key. Document retrieval again is driven by what we have termed "real communication". Interests are matched against the ontology embedded in the document yielding graded matches. Search and retrieval can thus be guided by the intuition of the individual and the context that he operates within. Both learning and problem solving will determine a major part of the professional's interest. A flexible indexing system as the one described will enable historic documents to be mobilised efficiently for issues at hand, both new and old. This is illustrated in figure 4.

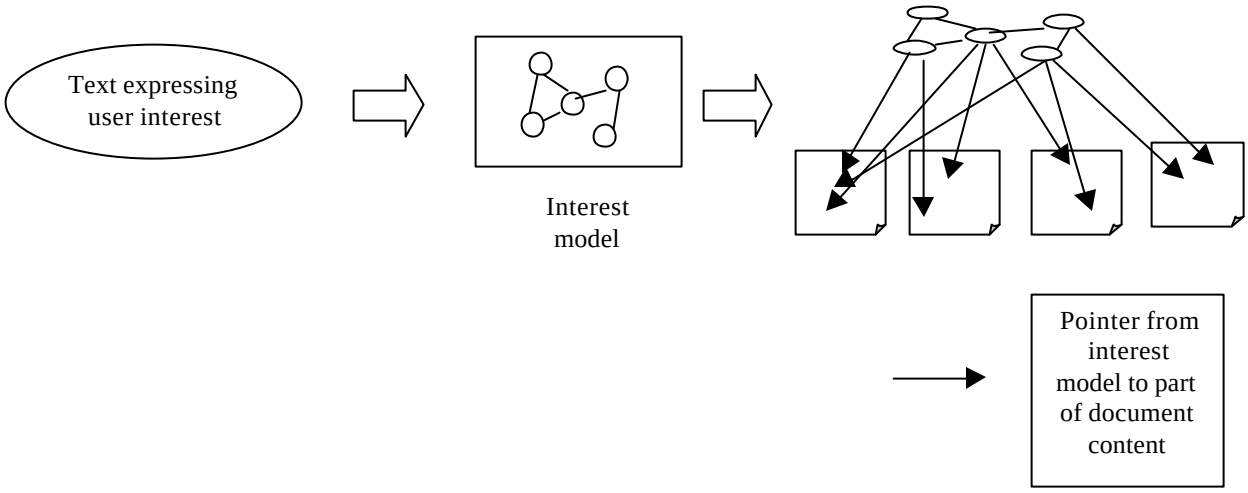


Figure 4. Using the interest model as a template for indexing

Analysing documents

Ontologies embedded in text will define its content. Central concepts constituting the core can be listed once the document is analysed. This summary function provides a simple, but important overview of the text found and can serve as a useful "super abstract". In the event that agents work on behalf of the user it is important that the analysis and match operation is made transparent. It is important that the agent conveys its findings and the match results in a way that enables the user to make decisions based on that. Explanations are central in that respect. The shared ontology between the analysed document and the interest model defines the rationale for why a document is important to read. The indexing system is non-binary. In other words it provides an analogue representation of the fit between interest and document. This has been exploited in order to show how the information is distributed across the document. A simple histogram will tell the user what part of the document is pertinent to the issues that he is interested in knowing more about (see figure 5).

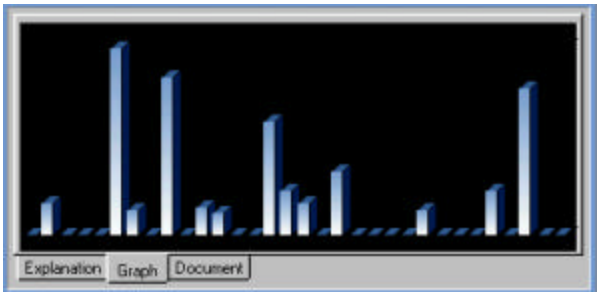


Figure 5 Histogram showing where the desired content in the document can be found and to what degree it is pertinent. Top of document is to the left. End of document is on the right hand side of the diagram. If a bar is selected and double-clicked the system will launch the paragraph represented by the bar.

Large documents do not necessarily yield large ontologies. Given a strict domain oriented focus a book might accommodate semantic models that are small, but extremely rich. This seems to be typical for textbooks that treat a subject matter in depth. The resulting ontology will then circulate around a core that stands out from less dominant aspects. This enables the system built in the TEKMO project to identify semantic links between objects that are seemingly independent. For this purpose simple search mechanisms can be applied.

Fusing and mining documents

In the outline above we have focused on bilateral communication using documents. Multilateral communication can be achieved when fusing documents. This implies that ontologies stemming from several different sources and authors can be superimposed on each other yielding a shared ontology. This type of ontology can support knowledge building. It supports text mining in a convincing manner exposing the dominating concepts (most shared aspects) as well as anomalies. By using the analysis techniques discussed in the previous paragraphs on single documents connections between output from seemingly disparate sources can be found. Such implications can expose rare, but strong co-variances as well as causal relationships. A case in point could be routine maintenance reports from different teams at different plants using the same type of equipment. Reports from one source may not necessarily expose causal relationships before they are combined with input from other sources. Bremdal et al has studied this application in more detail (Bremdal 99). Here we will point to a simple example that highlights the issue. The following list shows excerpts from different source text.

Text 1:	"the boiler was replaced after a spill of water on March 24, 1994"
Text 2:	"a leak was detected in the utility room at September 3rd, 1996"
Text 3:	"maintenance is done on a routine basis on pipes and seals in the boiler room"
Text 4:	"our prime maintenance goal is to avoid leaks"
Text 5:	"poor maintenance may cause spills"
Current issue:	"there is a large water spill in the boiler room"

When blending the current issue with the history and applying a fuzzy intersection algorithm on the fused ontology the system is able to draw the following connections between the main entities in the query:

Large water spill -> water -> boiler -> boiler room
Large water spill -> water -> utility room -> room -> boiler room
Large water spill -> spill -> utility room -> room -> boiler room
Large water spill -> spill -> boiler - > boiler room

This implies that the user query can be associated with the following:

Water suggests boiler in the boiler room.
Water suggests a utility room and room like the boiler room.
Spill suggests a utility room and room like the boiler room.
Spill suggests boiler in the boiler room.

If we compare with the historic material the inferences made by the system yields fairly good hints as to where to look for water spills and in association with what. Note that the lexical aspects, for instance the choice of words to use is of lesser importance. One example is the information given on a leak in the utility room. The system substitutes "leak" with "spill" and makes a semantic connection of importance.

This knowledge building mechanism generates benefits beyond summaries. It is very useful for business intelligence and knowledge refinement with respect to synthesis of experience. An example of the first is fusion of news articles and documents associated with a competitor and his customer. One article issued at some time may elaborate on some needs of a customer. Another might be a brief statement that a competitor has established a contract with a customer. An analysis of the competitor's home page on the WWW will probably disclose news about their interests, capabilities and fragments of their strategy. When the three items are fused the system might point out interesting causal relationships between the need of the customer and the capabilities of the competitor that explains the recent contract agreement. Despite inferior inferencing and information extraction capacities compared to that of a human it is possible for the system to process much larger quantities of information in a shorter time and maintain established relationships without forgetting these. This might yield a very good support for the human part that will generate an extensive overview and receive guidance about where to look further for increased details. The learning effect can thus be substantial and have a profound effect on future business operations.

Example

In the following we will illustrate the Mark II version of the system generated in the TEKMO project. Although experimental it has been fielded in different versions for various purposes. Below we have illustrated this by a simple use-case for subscribing to knowledge pertinent to a problem solving situation i.e. an engineering project.

A. The user defines the objective of his work and specifies the background knowledge for this. In conjunction with this he adds other ambitions of importance. For this purpose he uses a simple word processor editor. Basic behaviour follows the recommendation for sound project initiation by paraphrasing and reiterating on the project objectives (Figure 6a)

B. The specification is loaded into a personal agent that uses this information to serve the user. The agent rebuts on the information to expose its understanding of the mission. This paraphrase is made in terms of tokens that summarise the content of the interest model that the agent has built for the user. (Figure 6b)

C. The user specifies the regime that the agent should search and monitor. Included here is the server area where all agents are monitored. This is to serve the agent broker functionality and enable notification of other people's interest. The agent is also given report and notification instructions. The targets could include report format preferences, instructions with respect to information sharing with other agents and thus other people that are represented on the net (Figure 7)

D. Once active the agent may monitor the work in other projects at other ends of the company. It may search the Internet and collect ideas and input relevant for the objective specified. Only reports that are semantically relevant with the specification will be reported. The user is spared the effort of sorting out things. The agent will be the active part in the Intranet environment and maintain the awareness and increase the perception capability of both active and passive users. (Figure 8)

E. Once interesting things have been found and reported it can be indexed and stored for easy future reference. This makes the Intranet a real Corporate Memory. The Corporate memory will support intuitive queries, knowledge mapping and document fusion (Bremdal 99).

F. Documents related to the same context or interest can be fused in order to generate a data model encompassing information gathered from more than thousand pages[1]. This model can be queried and explored in ways similar to an object-oriented database.

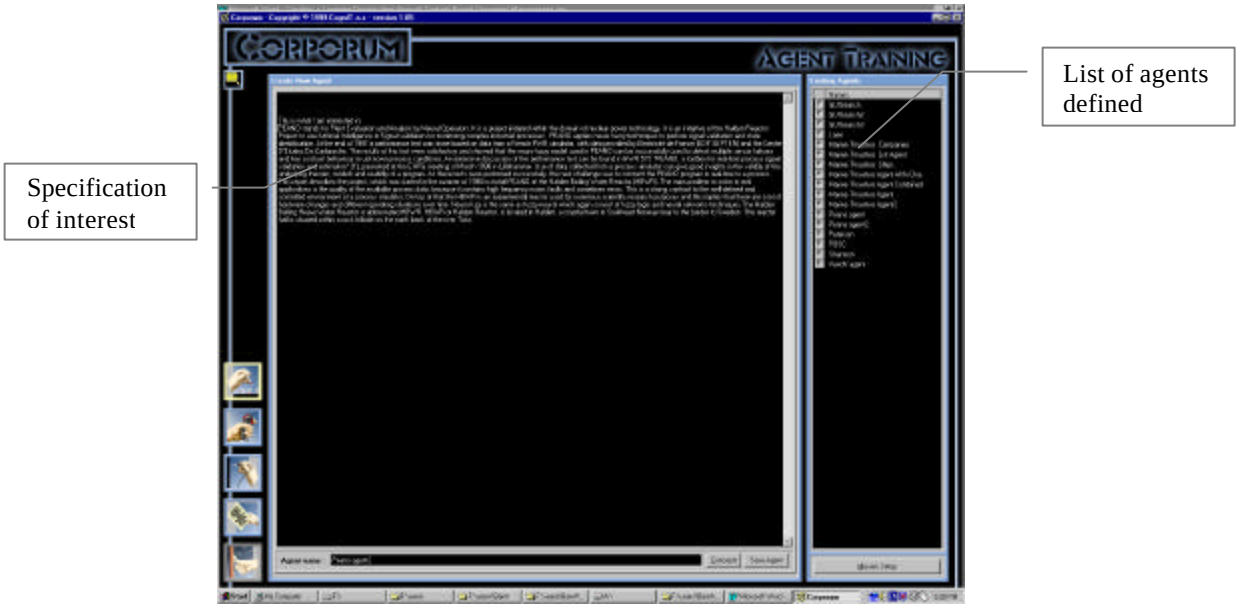


Figure 6a Defining scope of interest and creating agents. a: User enters text

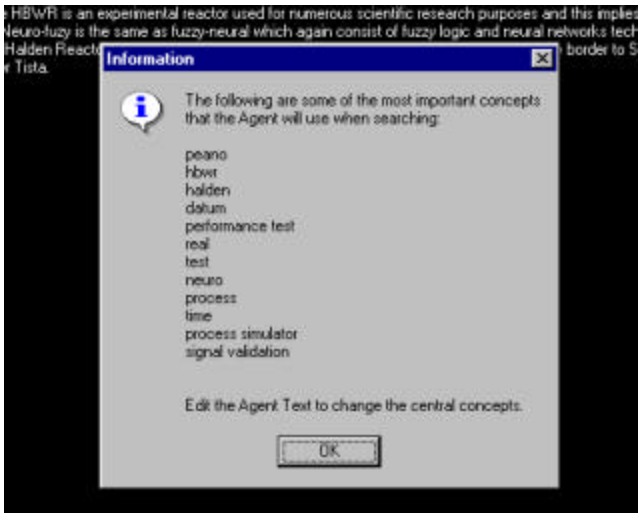
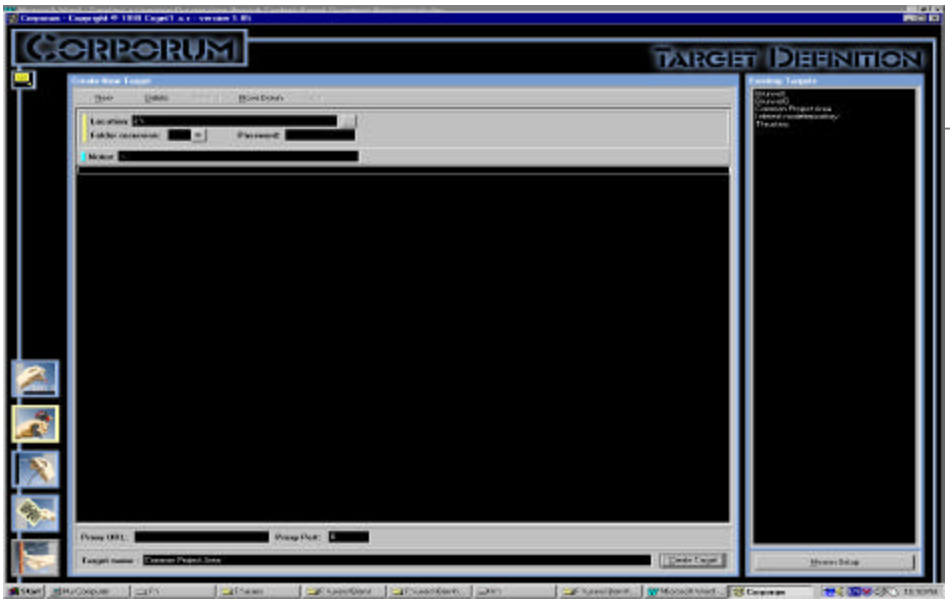


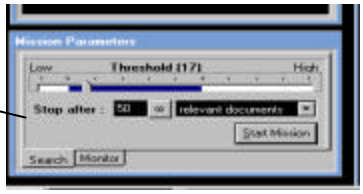
Figure 6b Defining scope of interest and creating agents. b: Training yields a feedback in the form of tokens that convey the agent's comprehension of the user focus.

[1] This is basically an issue of server capacity. There is, in principle, no limit with respect to number of pages. However practical tests carried out so far with ordinary NT servers suggest practical limits in the range of 800 - 1200 A-4 pages to one data model. But optimisiation techiques that enable normalization and reduction of the data model may multiply existing capacity in the magnitude magnitude of 10^{2-3} pages.



Target areas defined can include wide search regime on the WWW specific information exchange areas such as a dedicated folder or database in the Intranet.

Here we have a regime for interest broking as well as a common information space for project pertinent material



Defining agent parameters related to reporting: High threshold means very stringent document evaluation filtering

Figure 7 Defining target areas and operating parameters for agents

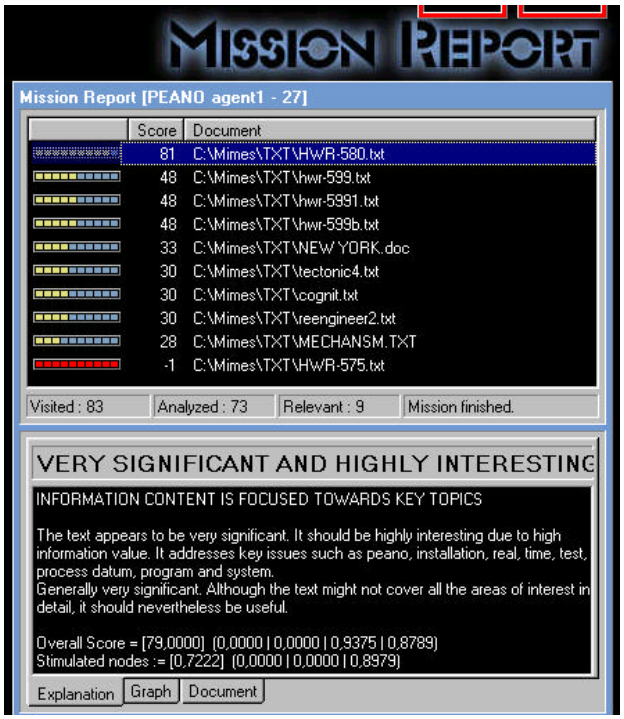


Figure 8 Agent list findings according to assumed interest. The findings are categorised and explained.

Final discussion and conclusion

In the TEKMO project we have addressed the needs of the Learning Organisation. We have developed a content based document concept based on agents that support a practical approach to achieve the objectives of the Learning Organisation. The system created exploits standard and novel document practices and the infrastructure that computer networks provide. The working hypothesis of the project has been that collections of documents made available through the Internet/Intranet constitute important knowledge repositories. The knowledge albeit partial and sometimes incomplete, can be captured and externalised through the use of ontological founded processes. In this manner knowledge building, knowledge mapping and knowledge sharing can be achieved. However, all of these require persistent involvement of people representing different professions on the net. The project has also treated this. The key has been the use of software agents that is driven by user interests.

The results achieved this far prove that it is possible to both extract and index knowledge contained in a document. The use of ontologies makes it possible to match different specifications of interests. It supports efficient data grabbing, subscription to pertinent domain information, search and document retrieval in an intelligent manner which focus at the knowledge content rather than an ad hoc categorisation such as a set of keywords or a standard archive tree. This provides for a personalised form of information service. The technicalities surrounding this have been highlighted in our presentation. All of these capabilities have been accommodated in the software agents that operate freely in a network environment and act strictly on behalf of its user.

The Intranet links people from remote parts of an organisation regardless of traditional barriers such as geographic distance, department divisions, different management policies of subsidiaries and time differences. We support the claim that the Intranet promises significant support for knowledge management functions. However, the Intranet is inherently static. It provides the capacity for substantial interaction between people, but does not automatically generate value-focused interaction. Fruitful collaboration on the net has been demonstrated, but it is required that people are familiar with the kind of socialisation and communication that the net can provide. This does not need to be true of all organisations, unless measures are taken on the management side to actively encourage this through proper incitements. Data overload represents a challenge for a very active organisation. Means to capture the essence of an interaction is imperative. All of this will have to cover a broad spectrum of initiatives.

Based on our experience with initial field testing of CORPORUM in the TEKMO project we believe that agents can serve as interest brokers. In that manner they can also help bring people together on the net. They can help to sort the meaningful from the meaningless. Contextual knowledge can be captured and channelled to the desk where it will be instantly put to use. Since agents are able to determine knowledge profiles and focus of interest in written material they can establish what overlapping competence and interests authors and active net participants have. Based on this knowledge they may initiate professional relationships by actively prompting people for contributions in terms of advice, invitations and support. They may set up shared information buffers for particular subjects and relay results of interactions between a few participants to a number of observers. In this way we believe that it is possible for CORPORUM to actively support an adhocractic organisational structure that expand and contract with the presence and absence of new problems and infused dynamics. This structure can be built on top of a more efficiency oriented kind of organisational design. Each member of the organisation may then perform and communicate in a balanced manner. The real effect will be, in our opinion, an organisation that achieves a more liberal flow of knowledge. It is an organisation that will prosper on more effective distribution, reuse and fusion of written material. Team approaches are possible across very large and scattered organisations. The concept that we have described may represent a new and important platform for creating a Learning Organisation.

As a final word we would like to emphasise that the focus in TEKMO is basically technical. That does not imply that we believe that technical solutions alone can solve all the practical issues related to the introduction of the Learning Organisation. Human aspects such as motivation, skill, attitudes and values are all very important. In our discourse we have visited a few through our discussion on publication skills, socialisation and the use of open discussions to make reasoning transparent. Much new thinking and significant effort will have to be added in order to approach the new media that we have prescribed. In order to demonstrate our concept in practical work surroundings and in order to improve our understanding of the human aspects related to the concept a substantial part of our future R&D will be directed towards extensive field testing. A prime objective is to gauge the impact of CORPORUM on both practical problem solving and organisational learning.

Acknowledgements

This work has been financed in part by support from The Research Council of Norway, Norges Forskningsråd (NFR).

References

(Aamodt 91) Agnar Aamodt. A Knowledge-Intensive, Integrated Approach to Problem-Solving and Sustained Learning. PhD Dissertation NTH 1991:7. 1991

(Argyris 91) Chris Argyris. Teaching Smart People How to Learn. Harvard Business Review. May -June 1991

(Bremdal 99) Bernt A. Bremdal. An Agent-Based Corporate Memory for Maintenance and Safety Engineering. HRP Meeting, Loen, Norway. May 1999

(Bremdal 98) Bernt A. Bremdal. Communicating Tacit Knowledge. Governing principles for externalisation. CognIT Research Report. 1998

(Davenport 98) Thomas H. Davenport & Laurence Prusak. Working Knowledge: How Organizations Manage What They Know. Harvard Business School Press. 1998

(Drucker 88) Peter F. Drucker: The Coming of the New Organization. Harvard Business Review. Jan-Feb 1988.

(de Hoog 94) Robert de Hoog, Rob Martil, Bob Wielenga, Rob Taylor, Clive Bright, Walter van de Velde. The Common KADS model set. Document Id. KADS-II/M1/DM1.1b/UvA/018/6.0/FINAL. University of Amsterdam 1994.

(Fielding 97) Roy T. Fielding and Gail Kaiser. Collaborative Work: The Apache HTTP Server project IEEE Internet Computing. July/August, 1997

(Feldman 98) Ronen Feldman. Text Mining: Theory and Practice. 4th World Congress on Expert Systems - Application of Advanced Information Technologies. ITESM Mexico City Campus. Mexico City. March 16-20, 1998.

(Fiol 85) C. Marlene Fiol and Marjorie A. Lyles. Organizational Learning Academy of Management Review. October 1985

(Gannon 98) Andrew Gannon. Knowledge Management at Hewlett-Packard. Knowledge Management Vol 1, # 3 Dec/Jan 98

(Garvin 93) David A. Garvin. Building a Learning Organisation Harvard Business Review. July-August 1993

(Genesereth 92) Genesereth and Fikes. Knowledge Interchange Format. Version 3.0 of Reference Manual Logic Group, Report Logic-92-1, Computer Science Department, Stanford University. 1992

(Gruber 93) Thomas Gruber. Towards Principles for the Design of Ontologies Used for Knowledge Sharing. International Workshop on Formal Ontology, Padova, Italy. March 1993

(Hamon 96) Jocelyne Hamon, Gildas Leroy, Christophe Spaggiari, Bernt A. Bremdal, Rolf-Einar Grini, Svein Nilsen. ACCES - A Tool for Capitalising and Utlilising Knowledge. HRP Meeting, Loen, Norway. 1996

(Heygate 94) Richard Heygate. Being Intelligent about "intelligent" technology The McKinsey Quarterly Number 4. 1994

(Kruisinga 98) Eelco Kruizinga, Gertjan van Heijst, Rob van der Spek. Knowledge Infrastructures and Intranets: Re-engineering for Learning Part 2.Knowledge Management. 1999

(Levitt 89) Barbara Levitt, James G. March. Organizational Learning American Review of Sociology, Vol 14. 1988

(Mikkelsen 96) J.O.Mikkelsen, C.B. Stabell, R.Sinding-Larsen: Knowledge-Based Advice & Critique Support for Play and Prospect Evaluation. In Artificial Intelligence in the Petroleum Industry Vol. 2 (Ed: Braunschweig and Bremdal), Éditions Technip, Paris 1996.

(Mintzberg 83) Henry Mintzberg. Structures in Fives: Designing Effective Organizations Prentice-Hall International Editions. 1983

(Rittel 72) H. Rittel. On the Planning Crisis Bedriftsøkonomen No. 8, Oslo. 1972

(Schreiber 98) A.Th. Schreiber, J.M. Akkermans, A.A. Anjewierden, R. De Hoog, W.Van de Velde, B.J. Wielenga. Common KADS - Engineering of Knowledge: The CommonKADS Methodology [version 5.0) Department of Social Science Informatics, University of Amsterdam . 1998

(Senge 90) Peter M. Senge. The Fifth Discipline: The Art & Practices of the Learning Organisation Doubleday. 1990

(Shannon 48) Claude E. Shannon. A Mathematical Theory of Communication Reprint from the Bell System Technical Journal Vol. 27, pp. 379-423, 623-656, July, October 1948.

(Simon 76) Herbert A. Simon. Administrative Behavior: A Study of Decision-Making Processes in Administrative Organizations (3rd Edition).The Free Press. 1976

(Speh 98) Marcus Speh. KMunity Building at Shell. Knowledge Management Vol 1, # 3. Dec/Jan 98

(Stata 89) Roy Stata. Organisational Learning - The Key to Management Innovation Sloan Management Review. Spring 1989

(Stokke 94) Per S. Stokke and Bernt A. Bremdal. Information Refineries and the manufacturing of Industrial and Corporate Knowledge Proceed. International Seminar on the Management of Industrial and Corporate Knowledge, ISMICK 94, Compiégne, France. October 1994

(Swartout 99) William Swartout and Austin Tate. Guest Editor's Introduction: Ontologies IEEE Intelligent Systems and Their Applications. January/February 1999

(Wilson 98) Owen Wilson and Corinne Sellens. Walk Before You Can Run. Knowledge Management Vol 1. Nov. 1998.