UNIVERSITY OF Southampton

University of Southampton Research Repository
ePrints Soton

Copyright © and Moral Rights for this thesis are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given e.g.

AUTHOR (year of submission) "Full thesis title", University of Southampton, name of the University School or Department, PhD Thesis, pagination

http://eprints.soton.ac.uk

AN ONTOLOGY-BASED PLATFORM TO SUPPORT ORGANISATIONAL INNOVATION NETWORKS

By

Richard Michael Beales

A thesis submitted for the degree of Doctor of Philosophy

School of Electronics and Computer Science
University of Southampton
United Kingdom

December, 2004

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE & MATHEMATICS

SCHOOL OF ELECTRONICS & COMPUTER SCIENCE

Doctor of Philosophy

AN ONTOLOGY-BASED PLATFORM TO SUPPORT ORGANISATIONAL INNOVATION NETWORKS

by Richard Michael Beales

Successful organisations innovate. Innovation in operational and business practice is essential in any organisation wishing to increase efficiency and profitability and adapt rapidly to changing external conditions, while changing fashions, the threat of terrorism, environmental concerns and diminishing natural resources all encourage demand for products and services that are not just new, but *different*. Despite this, the potential market for tools that can support innovation has largely been ignored by business software providers.

To understand the means by which innovation *can* be supported, this thesis first outlines the close relationship between innovation and knowledge, examines mainstream organisational knowledge management processes and considers their limitations when applied to innovation-led organisations. An alternative, community-based knowledge management model is introduced that better suits these organisations. Existing technologies that address some element of this model are surveyed and their limitations identified. These findings inform the specification of a flexible, holistic innovation support architecture, Aware, based on a unified organisational conceptualisation, or ontology.

Aware is an active platform, implemented using the JESS expert system shell, which maintains a dynamic organisational evaluation that supports a range of innovation support services. These include bulletin delivery, event management, social network identification and expert finder capability. It is shown that this single platform addresses all of the key environmental requirements of innovation, while providing functionality and a degree of integration that cannot be achieved with traditional multi-system approaches. The platform is demonstrated using a kiosk client application for deployment within a single innovation-led organisation and further tested by exploration of an alternative usage scenario within the organic farming community.

Contents

List of Figures

DECLARATION OF AUTHORSHIP

I, Richard Beales, declare that the thesis entitled “AN ONTOLOGY-BASED PLATFORM TO SUPPORT ORGANISATIONAL INNOVATION NETWORKS”, and the work presented in it are my own. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.
- none of this work has been published before submission.

Signed:

Date:

Acknowledgements

I gratefully acknowledge the patient support and guidance of my supervisors, Nigel Shadbolt and David De Roure. Thanks are also due to Ernest Friedman-Hill at Sandia National Laboratories for his development of the JESS rule engine on which this work depends.

This work was supported by the Advanced Knowledge Technologies (AKT) Interdisciplinary Research Collaboration (IRC). The AKT IRC is sponsored by the UK Engineering and Physical Sciences Research Council under grant number GR/N15764/01 and comprises the Universities of Aberdeen, Edinburgh, Sheffield, Southampton and the Open University.

Chapter 1

Introduction

Successful organisations innovate. Innovation in operational and business practice is essential in any organisation wishing to increase efficiency and profitability and adapt rapidly to changing external conditions; it is one of the primary forces of economic improvement. Social innovation capital, the ability of an organisation to generate innovation through collaboration, is duly recognised as a contributor to its intellectual capital and as such, an important predictor of its future performance [Power98]. Many organisations, though, exist *to* innovate. Research universities, commercial scientific laboratories, high-technology companies and producers of luxury consumer goods are in a constant race to develop new products, new technologies and new techniques. Changing fashions, the threat of terrorism, environmental concerns and diminishing natural resources have all encouraged demand for products and services that are not just new, but *different*.

Surprisingly, commercial business software developers have largely ignored innovation and concentrated instead on products that optimise well established tasks and processes. This is not to say that there is no software support for novel or imaginative work. Indeed, some of the biggest names in the software industry – Adobe, Autodesk, Macromedia, Pinnacle and others – are dedicated to producing tools that enable the realisation of creative ideas. Beyond the creative industries, computer simulation in science and manufacturing allows new processes or materials to be trialled virtually, at minimum cost and risk. Computer aided manufacturing (CAM) processes, such as CNC (Computer Numeric

Control) milling, MELATO[1] (Metal Laminated Tooling) and ThermoJet[2] 3D printing, allow one-off physical prototypes to be produced quickly and cost-effectively. The growth of digital publishing and the Internet has meant that news of new products or services can reach a far wider audience of potential investors and customers than ever before.

What is notably absent, though, is software that addresses the broader processes of innovation-led organisations. We are constantly reminded of the need to disseminate best practice and there is no shortage of software that claims to achieve this, but where is the software to encourage *new* practice? The suggestion that innovation can somehow be managed or stimulated electronically might sound rather improbable, if not a little sinister. However, innovation is known to be heavily dependent on the sharing of knowledge between individuals with different experiences, interests and expertise. If systems can be put in place that break down organisational barriers and encourage the free exchange of knowledge throughout the organisation, an environment more conducive to innovation is created.

In this thesis, an understanding is formed of the essential role of individual knowledge to innovation and, more specifically, of the importance of the exchange of this individual knowledge through informal face-to-face discussions rather than electronic knowledge bases. It is found this free, informal knowledge exchange depends on the existence of dynamic, rapidly evolving active physical communities between which there is strong awareness and effective communication. Each of these three challenges, awareness, community and communication, is, individually, already served by a multitude of electronic tools, ranging in sophistication from simple mobile devices that alert the user when a colleague is nearby, to kiosk-based community bulletin boards and professional

[1] www.fraunhofer.de/english/press/pi/pi2002/index_11_t5.html

[2] www.3dsystems.com/products/multijet/thermojet/index.asp

network exploration applications for the desktop PC. Yet, though each depends on some access to underlying organisational concepts and relationships, these independent, standalone devices and applications have no common underlying information model. While much information is duplicated between them, when one has access to information that another lacks, this cannot be shared. The user is inconvenienced unnecessarily, being required to manage multiple copies of the same information. Inconsistencies emerge both between the information models themselves and between the way that different applications interpret the same information. Most importantly, it is argued, these individual tools do little to foster interaction between different communities, instead tending to reinforce existing boundaries and further isolate their members.

To counter the limitations experienced with existing tools, it is proposed that a single, carefully constructed organisational conceptualisation serve as the basis of a unified approach to innovation support. The solution developed depends on two areas pioneered by the Artificial Intelligence community: ontologies and inference engines. At its simplest, an ontology is a hierarchical conceptualisation, or model, of a domain; it is a class structure, but one in which the relationships *between* classes are as important as the classes themselves. Inference engines act on combinations of facts and relationships – such as those represented in an ontology – to perform actions specified by pre-loaded rules.

These two technologies are combined to form the basis of an innovation support platform. On top of an active organisational conceptualisation is situated a palette of awareness, communication and communication functions, including support for electronic bulletin boards, community formation and management tools and expert and common-interest visualisations. These can be accessed individually, by simple single-function mobile devices, or daisy-chained together to support sophisticated kiosk and desktop PC applications; example mobile and kiosk clients are described in the thesis. In each case, the underlying information model that is used is the same and the behaviour exhibited to the user consistent.

1.1 Contributions

The work contained in this thesis makes two significant contributions to the fields of Computer Supported Cooperative Work and Knowledge Management. Firstly, a common organisational ontolog, and an actively maintained network of numeric evaluations of concept relationships within the ontology are presented as the basis of a single platform to support three key requisites of innovation: awareness, community and communication. This is feasible because each of these requisites depends largely on an understanding of the relationships between different personal interests and involvements within an organisation. If it is determined, for example, that two individuals share an interest in a particular subject, that piece of information can be used by a wearable electronic ice-breaker device to introduce them when they are physically proximate; equally, if one of the individuals contributes their interest to a project group, it can suggest that details of that group should be communicated to the other person.

In contrast to existing, piecemeal organisational IT strategies, in which awareness, community and communication are addressed individually by entirely independent applications, this unified approach offers substantial benefits. Rather than imposing artificial distinctions between awareness, community and communications support, it allows composite applications to be developed that combine elements of each to better reflect the way that individuals naturally manage their social and collaborative networks; task-oriented tools such as a meeting room assistant or foyer information kiosk can support a seamless blend of awareness, community and communication functions tailored to the situation in which they are deployed.

Relative to a discrete approach, the sophistication of support afforded to the individual requisites in this unified platform is also significantly greater. Because evaluations of the platform's information model are maintained continually, relationships that would previously only have been uncovered by an explicit query by a user can instead be brought to their attention automatically, immediately they are formed. They are also available to be used by lightweight clients that, ordinarily, would not justify the computing resources necessary to evaluate these relationships on demand, and in applications that would otherwise demand a frequency of evaluation of relationships such that they could not viably be implemented, as in the delivery of message bulletins according to a specified target interest.

The flexible alternative that the platform offers to absolute addressing and delivery of message bulletins is the second key contribution of the thesis. One of the most common challenges faced in all organisations is getting a piece of information to the ‘right people’; in innovation-led organisations, where individual employees’ interests and involvements change continuously, the problem is particularly acute. At present, a variety of electronic communication methods are used to disseminate information within organisations. However, whether a message is propagated via a personal email list maintained within a desktop email client, a centrally managed, formal mailing list, an intranet web forum or a mobile phone text message group, its author is required to know, at the time the message is despatched, the specific individuals or groups of individuals to whom it should be delivered.

Here, a delivery mechanism is presented in which messages are addressed not to specific individuals, mailing lists or display devices, but instead to a *description* of their target audience that is evaluated dynamically by the delivery platform. This description, or target profile, allows an individual to scope the delivery of their message to those with interests similar or relevant to their own, to the memberships of project groups specified directly or by example, and to attendees of events, forthcoming or historic. The target profile also allows the time period during which the message should be visible to be restricted and, though absolute time limits are supported, again the need to provide specific addressing information is obviated by the ability to define the display period relative to the lifetime of a project group or to a future event, regardless of whether that event has yet been scheduled.

As target profiles can be re-evaluated in response to changes to the system's conceptualisation of the organisation, messages can automatically be made available to others to whom they later become relevant, provided their specified display period has not expired. Furthermore, as 'delivery' of a message involves the population of the target profile with lists of virtual destinations that should have access to the message at the present moment, rather than replication and transmission of the message content itself, its author is free to edit the message, or retarget or retract it, at will, long after it has first been posted. Despite the power of this message targeting technique, a sample client application presented in the thesis demonstrates the ease with which a target profile can be constructed as an unobtrusive step in the message authoring process.

1.2 Document Structure

Chapter 2 identifies the close relationship between innovation and the free exchange of knowledge. The limited success that has met deployments of mainstream, information-based Knowledge Management methodology in innovation-led environments is discussed and motivates consideration of an alternative, *community*-based approach to Knowledge Management, in which individuals, rather than systems, are pre-eminent.

Though this treatment presents the first coordinated software approach to supporting innovation, a range of tools and devices has already been developed that assist aspects of community-oriented work. Chapter 3 summarises a representative selection of inter-personal network visualisations, information exchange platforms and inter-personal awareness aids.

In Chapter 4, the awareness, community and communication requirements of innovation-led organisations are considered. The underlying importance to these concepts of individual *interest* is identified. Interests are then situated in a wider organisational context as part of an ontology developed to capture an actionable organisational model. The architecture of a unified innovation support platform that will host this ontology, Aware, is introduced.

Chapter 5 details significant aspects of the Aware platform's implementation. The rules that define the behaviour of its active domain conceptualisation are examined; particular consideration is given to mechanisms that maintain the network of evaluated topic, interest and interest group relationships that are fundamental to most of Aware's services. The operation of two such services, which identify subject expertise and social networks, is described.

While it would be meaningless to present a definitive client application -- both the platforms used and the compositions of services that they present will depend upon the environment in which the platform is deployed – a usage scenario provides the best opportunity to situate the individual platform services in an operational context. Chapter 6, therefore, considers a client application for use within a technical research organisation. A number of components are defined and the implementation of one, a kiosk client for semi-public spaces, is described.

In Chapter 7, the functionality provided by Aware is evaluated against the requirements identified in Chapter 4. The flexibility of the platform is tested by considering its suitability to support innovation in an environment that contrasts sharply with that discussed in the previous chapter: organic farming. Finally the active, self-evaluating conceptualisation on which Aware is based is contrasted with a more traditional query-led architecture.

Chapter 8 concludes the thesis with a summary of the work presented and outlines some opportunities for future work, suggesting technical enhancements to the platform itself and briefly exploring some alternative client applications that it might support.

Chapter 2

Innovation and Knowledge

"*Innovation is knowledge intensive. Both are interrelated and have to be addressed simultaneously*" UNESCO, 2003.

In this chapter, we will outline the close relationship between innovation and knowledge, examine mainstream IT-based organisational knowledge management (KM) processes and introduce ontologies as the structured representations of information on which these IT-based KM approaches are based. We will then consider the limitations of mainstream KM when applied to innovation-led organisation and explore an alternative, community-based KM model that better suits these organisations.

2.1 Innovation and Knowledge

Most commentators agree that innovation is a knowledge intensive process [Choo98] [Castells96]; it demands the straightforward application of knowledge – Drucker, for example, writes "*if we apply knowledge to tasks that are new and different, we call it innovation*" [Drucker92, p26] – but is also dependent on the *exchange* of knowledge between different groups or individuals, Arundel noting that "*modern innovation theories stress the diffusion of knowledge among many different actors*" [Arundel02]. The importance of knowledge to the modern economy first gained academic prominence in the 1960s. In that decade Franz Machlup coined the phrase "*knowledge industries*" [Machlup62], Peter Drucker observed that "*even the small business today consists increasingly of people who apply knowledge, rather than manual skill and muscle, to work*" [Nickols04] and

Alfred Marshall stated that "*knowledge is our most powerful engine of production*" [Venters01]. It would, however, be another 30 years before mainstream business would give serious consideration to knowledge as a critical organisational resource and 'Knowledge Management' (KM) entered the business lexicon, perhaps reflecting a belief that anything of commercial value must be manageable.

Though the term is liberally applied, within declared corporate KM strategies, two key objectives can be identified. One is to improve dissemination of best practice [OLeary00] and acquired knowledge within the organisation, a desire echoed by that often-repeated phrase attributed to former Hewlett Packard CEO, Lew Platt: "*If only HP knew what HP knows, we would be three times more productive*". One of the most widely known best practice initiatives is Ford Motor Company's Best Practice Replication (BPR) process [Dixon00]. BPR began in 1995 as a register of 45 best practices established during exchange visits between engineers at US and German factories. By 2000 the register contained over 600 ideas, with contributions from each of Ford's 37 international vehicle plants. On average, 40% of plant task practices are drawn from BPR and the operations manager of each plant is sent up to 8 further suggestions every week, each spelling out the savings in man hours, materials or energy it afforded at a sister plant. Staff are motivated to support the BPR system by the requirement – known as 'the Task' – that each plant achieve a 5% yearly productivity increase. Davenport notes a similar system in operation at Hughes Electronics [Davenport97], where a database of "*best process reengineering practices*" has been established. Another Ford initiative, the TGRW, or 'Things Gone Right / Wrong' [OLeary98] [Stewart97] knowledge base, captures information about events that facilitate or hinder task accomplishment; O'Leary notes, though, that employees are less keen to submit 'TGW' reports as they do not want to be associated with failures. At Rolls Royce, the SPEDE methodology [Hammersley99] has been employed to combine internal BPR with awareness of industry-wide best practices.

The second significant focus of corporate KM strategies is the safeguarding and exploitation of organisational knowledge as a commercial resource. Organisations increasingly recognise knowledge as one of their most important assets; an Ernst & Young survey of British entrepreneurs conducted in 2000 found 15% placed knowledge as *the* most important organisational asset, with a further 65% identifying it as one of the most important [Design00]. As this knowledge largely resides in the heads of employees, there is concern that organisations are vulnerable to 'knowledge attrition' – knowledge, literally, "*walking out of the door*" [Holloway99] [Swan00]. Newcombe [Newcombe99] quotes Delphi Group analyst Mark Tucker's observation that "*companies can no longer rely on their knowledge resources to stay*" and Becerra, taking as a base point Reimus's [Reimus97] estimation of a 10% annual employee turnover in IT consultancies suggests that, in a growing organisation, around 30% "*of the firm's knowledge workers lack sufficient depth in organizational-specific knowledge*" [Becerra99].

Organisations that are downsizing also employ KM to capture workers' knowledge before they leave. Santosus reports that as their B-2 stealth bomber production ended, defence contractor Northrup Grumman were faced with the loss of the expert knowledge that would be required to support and maintain the aircraft throughout its life: "*As the B-2 program was winding down and engineers with 20 or so years of experience were leaving,* (they) *established a 10-person KM team to identify subject matter experts and capture the content of their brain cells*" [Santosus01]. A number of off-the-shelf KM software packages are also marketed as protection for organisational knowledge from the effects of workforce churn or reduction: Epistemics[1], commercial developers of the PC-PACK Knowledge Acquisition toolkit [Milton99], suggest this product offers "*a reduced risk of sudden loss of expertise, particularly in areas where there is a reliance on just one or two experts*", VelocityStorm Software's Knowledgebase

[1] www.epistemics.co.uk

System[2] promises to "*capture employee knowledge so it doesn't walk out of the door when they do*" and Transversal Intrafaq[3] is advertised as alleviating "*the risk and costs associated with expertise leaving your organisation*".

Inevitably, having captured this material, some organisations have exploited its commercial value by marketing it externally. Hibbard reports that Dow Chemicals have used SmartPatent Workbench, a patent search tool, to identify valuable Dow patents that can be licensed to other organisations or traded for goods or services [Hibbard97]; O'Hara [OHara01] envisages that future semantic web and agent technologies will enable the creation of inter-organisational knowledge marketplaces.

2.2 Limitations of traditional KM

Despite the undoubted benefits that many organisations have gained from the traditional, 'IT-based' approach to knowledge management, it has widely been suggested that it is not an effective means to support innovation. One concern is that it is often impossible to predict how current knowledge will be used by future innovators. Majchrzak et al. note that "*innovation, by definition, means the use of knowledge in unknown future contexts and thus simple searches of any repository are unlikely to yield innovative outcomes*" [Majchrzak01]. Malhotra [Malhotra02] and Lambe [Lambe02] also question the feasibility of storing large quantities of knowledge for future reuse, Malhotra observing that the underlying assumptions a system must make when determining what should be retrieved cannot be "*pre-programmed to detect an unpredictable future*" [Malhotra02].

It is suggested, however, that a more fundamental limitation on the usefulness of electronic knowledge repositories is the kinds of knowledge that

[2] www.velocitystorm.com

[3] www.transversal.com

they are able to capture. Typically, emphasis is placed on their inability to handle something called 'tacit knowledge' [Nabeth01][Scarbrough01][Swan99];. Scarborough, for example, claims "*a particular weakness of the IT-based approach to KM… is the difficulty of codifying the tacit and highly distributed knowledge involved in the* (innovation) *process"*. In the following section, we will examine this claim, exploring the different classes of knowledge and considering their implications for innovation-oriented KM.

2.2.1 Classes of Knowledge

As we have discussed, it is suggested that IT-based KM initiatives cannot adequately support innovation in part because of their inability to exploit something called 'tacit knowledge'. Tacit knowledge was identified in the 1960s by chemist-turned-philosopher Michael Polanyi, who observed, simply, that "*we know more than we can tell*" [Polanyi66, p.208]. KM practitioners later divided knowledge into two distinct, mutually exclusive classes. Tacit knowledge, it was argued, formed one of these classes; knowledge that was not tacit was to be known as *explicit* knowledge. Tacit knowledge would include "*subjective insights, intuitions and hunches*", the collective 'expertise' of an organisation [Lutters00]. It was held to be "*highly personal and hard to formalise*" [Nonaka91], "*not codified, not communicated in a 'language'*" and "*manifest only in its application*" [Seidler03]. *Explicit* knowledge, on the other hand, was identified as that that can be "*expressed in words and numbers and can be easily communicated and shared in the form of hard data, scientific formulae, codified procedures or universal principles*" [Skyrme99].

While explicit knowledge would lend itself to direct manipulation by software tools and was responsive to IT-based KM techniques, it was claimed that tacit knowledge "*will typically be of* (most) *value to innovation processes*" [Swan00]. Indeed, Swan et al. suggested that tacit knowledge influences each step of the innovation process, from "*initial idea discovery and generation*", through “*ideas transformation and implementation*” and implementation testing, to market deployment and ongoing support thereafter. Nonaka and Takeuchi [Nonaka95] posited that tacit knowledge could be made explicit and vice-versa, through social interaction. They proposed a cyclical translation process, SECI, encapsulating four knowledge conversion mechanisms:

- tacit knowledge transfer through Socialisation
- tacit knowledge to explicit knowledge conversion through Externalisation
- generation of new explicit knowledge through Combination of existing explicit knowledge
- the acquisition by individuals of tacit knowledge through Internalisation of explicit knowledge

It is, however, difficult to reconcile the suggestion of a straightforward social mechanism for the conversion of tacit knowledge into explicit knowledge with the implication of Polanyi’s famous expression, that tacit knowledge ‘cannot be told’ [Hildreth02]. Moreover, Tsoukas questions Nonaka and Takeuchi’s claim that tacit knowledge is internalised explicit knowledge, arguing that it is instead “*the necessary component of all knowledge*” [Tsoukas96]. If Tsoukas is correct, how then is it that the stuff that Skyrme identifies as explicit knowledge *can* be codified effectively, when tacit knowledge cannot, or that computer ‘knowledge bases’ can store knowledge, when they do not contain anything that cannot be told? Though, as Whitley observes [Whitley00], these contradictions are often side-stepped in KM literature, they can be resolved by the introduction of a further knowledge class – ‘*implicit knowledge*’ – and reconsideration of the nature of explicit knowledge.

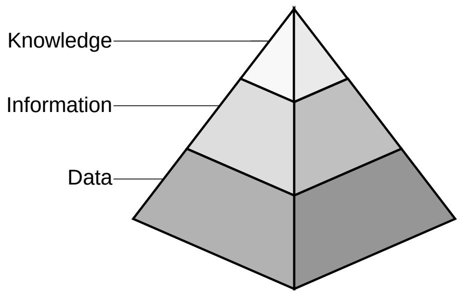


Figure 2.1 – Knowledge pyramid

Let us first tackle explicit knowledge. To do so, we need to examine the relationship between knowledge and information and data. As shown in Figure 2.1, the most common representation of knowledge places it atop a conceptual pyramid [Skyrme99], the foundation of which is *data* – raw symbols, sequences of numbers, characters on a page. Data is devoid of any context; it "*says nothing about its own importance or irrelevance*" [Davenport98]. *Information* is data organised, to give meaning; data with added value. It exerts influence (literally, it *gives shape*). Information, then, is 'more than' data. Knowledge, in turn, is *more than* information. It is information interpreted through human experience. You may be *informed* that a person is in two places at once, but you *know* this cannot be true.

This model suggests a unidirectional relationship between data, information and knowledge; information is derived from data and knowledge from information, with utility or 'value' gained by each transformation. Data, though, does not simply emerge from thin air. Tuomi [Tuomi99] illustrates this by considering temperature data obtained from a thermometer. He notes that the thermometer "*is created simultaneously with the possibility to observe temperature as data*". Not only is knowledge derived from this data, it enabled its existence. It is gained through interpretation of the recorded temperature, but it also informed the design of the thermometer – "*important aspects of knowledge* *are sedimented into* (its) *structure*" – and the manner in which it was used to obtain the temperature reading.

Information and data are not so much the basis of knowledge as the mechanism for its exchange. Knowledge "*becomes information once it is articulated*" [Alavi01]. The information is transmitted and stored in symbolic, or data form, as speech phonemes, visual glyphs, ASCII codes or some other form. From these symbols, information is reconstructed, which "*is converted to knowledge once it is processed in the mind of individuals*" [Alavi01].

Tuomi and Stenmark [Stenmark02] conclude, therefore, that explicit knowledge is not actually knowledge, but *information*. Further, they observe that for knowledge to be communicated intact between individuals, they must have relevant shared understanding; "*both the original articulator and the sensemaker need to have overlapping meaning structure*" [Tuomi99]. If they do not have this common ground, though they will both posses the same piece of information, the knowledge that they derive from it will not be the same. Recall the scientific formula, given by Skyrme as an example of explicit knowledge. The symbolic representation could be learned and recited by anyone, but the knowledge it would impart to a person is dependent on their scientific background. To an individual with no scientific understanding, it would remain a meaningless collection of symbols; uninterpretable data. An interested amateur might understand its literal meaning, its information content. Only an expert, though, would comprehend its scientific significance. This comprehension might be supported by background *cultural* knowledge. If, for example, the formula was written in 'old fashioned' script, it would be understood that it was likely not a recent scientific discovery.

Common ground minimises the effort required to communicate [Monk03]. If sufficient common ground is already in place, only the informational component of the knowledge – the written formula, in our example – must be transferred. The exchange need not take place in real-time as the information can be stored losslessly in a knowledge base for later retrieval. In effect, the knowledge exchange process outlined by Alavi and Leidner can be held in suspended animation, parked at the intermediate information stage. However, where pre-existing common ground cannot be assumed, where there is not sufficient of what Dahlbom and Mathiassen call ‘shared practice’ [Dahlbom95], establishment of common ground must accompany the transfer of information. It is possible to include in the stored information sufficient background that it can be understood by a non-expert, or phrase it in reference to another, related, subject area. This approach is commonly used in print publishing; some titles assume a high degree of prior knowledge, but others offer a ‘complete guide’ to a subject or provide ‘a primer on such-and-such for users of so-and-so’. Typically, though, non-expert treatments are only prepared when it has been ascertained that the subject concerned will be of widespread use or interest. The investment in time demanded of both the author of the information and its readers, who must digest the information and then determine its applicability, dictates that this is so[4].

Groups of individuals from different disciplines can, however, efficiently communicate relevant knowledge directly, through conversation. They are able to ‘feel out’ the common knowledge that they share, identify and correct misunderstandings and overcome confusion that might be caused by conflicting uses of terminology. This interactive negotiation of common background knowledge is known as *grounding*. Clark and Brennan [Clark91] identified eight constraints that influence the ease with which grounding can take place:

- Co-presence: do the individuals share the same physical environment?
- Visibility: can the individuals see each other?
- Audibility: can the individuals hear each other?

[4] Dahlbom and Mathiassen remind us that all publishing assumes some degree of prior common ground: *Think of what a cook book for the true novice would look like. Every recipe would begin: Turn on the light in the kitchen* [Dahlbom95]

- Contemporality: is there a delay between information being issued and received?
- Simultaneity: can individuals receive and issue information simultaneously?
- Sequentiality: does the interaction mechanism prevent an individual receiving information out of sequence?
- Reviewability: can information be reviewed after it has been exchanged?
- Revisability: can information be reviewed and revised before it is exchanged?

Oral face-to-face communication immediately satisfies the first six of these constraints and, if some form of note-taking is employed, reviewability is also achieved. This form of communication is, however, not revisable; in a real-time conversation phrases can only be retracted or refined after they have been spoken, by exchanging more information. Nonetheless, it is held to be the most effective means of grounding, partly because it affords the greatest awareness of paralinguistic and extra-linguistic cues such as tone of voice, hand and facial gestures and body posture that can connote uncertainty, puzzlement or misunderstanding [Clark96]. The effectiveness of face-to-face knowledge exchange does not, of course, preclude a role for electronic knowledge-bases or printed documents or books as sources of external knowledge to innovation groups, though it does emphasise that they should ideally be interpreted by a domain expert to whom the group has physical access. Fischer and Ostwald note, though, that the knowledge contained in the documents can be less important than the common ground that is attained through their discussion; their role as "*boundary objects*" is described, in which they act "*as reminders that trigger knowledge, or as conversation pieces that ground shared understanding*" [Fischer03].

The knowledge that grounding affords access to is not explicit knowledge. It is unlikely to have previously been codified or articulated, as it is a response to the particular common knowledge held by the participants of the grounding exchange. Nor, though, is it tacit knowledge; as Clark illustrates, it clearly *can* be 'told' and is exchanged rapidly through conversation, while tacit knowledge must be acquired gradually, over time, with growing experience, expertise or 'connoisseurship' [Polanyi62, p.54]. Rather, it is what some have labelled *implicit knowledge* [Nickols00] [Wilson02]; that is, simply, knowledge "*that can be articulated but hasn't*" [Nickols00]. Re-examining the SECI process, it is now apparent that Externalisation and Internalisation do not, as Nonaka and Takeuchi suggest, achieve translations between explicit knowledge and tacit knowledge – there is no mechanism by which this translation can take place – but between explicit knowledge and implicit knowledge.

This understanding of explicit, implicit and tacit knowledge has a number of implications for the design of KM systems to support innovation. Innovation depends on the "*integration of knowledge across disparate social communities*" [Swan00]. As formal knowledge-bases and document repositories will typically be directly usable only within the expert community by and for whom they were authored, it is important that external innovation groups using these resources can gain face-to-face access to community members that can re-interpret the published work for their particular application and expertise.

Tacit knowledge, we have learned, cannot be codified in knowledge-bases and documents, yet this knowledge, too, is important in innovation. Polanyi noted its role in problem identification, allowing experienced individuals to intuitively recognise issues that would not be apparent to non-experts referencing formal design manuals or knowledge bases [Polanyi66]. Tacit knowledge is also the basis of personal judgements of whether a product looks or feels 'right'; Miller, for example, refers to the tacit understanding of the sound that a luxury car door should make when it is closed [Miller99, p.94]. As tacit knowledge cannot be transferred quickly from one individual or expert group to another, it will be necessary not only to locate an individual who has this knowledge, but also to keep them engaged as the innovation progresses.

2.2.2 Towards a Community-based Approach

As noted by Scarbrough, "*it is clear that certain features of the innovation process pose a challenge for established approaches to KM*" [Scarbrough01]. Recognition of the value of person-to-person implicit and tacit knowledge exchange has led to proposals [Fischer01] [Groth01] [Nabeth01] [Scarbrough01] [Swan99] of an alternative, person-centred approach to KM that "*highlights the importance of relationships, shared understandings and attitudes to knowledge formation and sharing within innovation processes*" [Swan00]. Particular importance has been placed on providing support for the communities and collaboration networks through which informal knowledge exchange takes place [Carotenuto99] [Swan00] [Arundel02] – "*putting people in touch not with databases, but with other people*" [Lindgren021]. It is these communities and networks that we will explore in the next section.

2.3 Communities and Networks

Sociologists have proposed a range of interpersonal network theories to describe workplace relationships. Three – intensional networks, Communities of Interest and Communities of Practice – have been found to be of particular importance to innovation-led organisations.

2.3.1 Communities of Practice

The Community of Practice (CoP) was probably the first interpersonal network model to attract widespread interest from managers and business theoreticians. At its simplest, a CoP is a network of mutual improvement and support, formed by a group of people who have worked together over a period of time. It is, as the name suggests, a group of *practitioners*, with a strong emphasis placed on hands-on problem solving experience. The CoP concept is as old as civilisation. Notable historical CoPs include craftsmens' corporations in ancient Greece, mediaeval artisans' guilds and journeymens' Compagnonnages in 19th Century France. Current interest in CoPs, though, stems from a study conducted by Julian Orr [Orr96] in the late 1980s. Orr examined relationships between photocopier service technicians at Xerox. When a photocopier failed, it would display an error code. A technician was called. Each technician was provided with manuals that detailed a repair procedure for each error code. And, as far as Xerox managers were concerned, that was how the copiers got fixed. In reality the displayed code often did not correspond with the real error, or the manual stated that a particular error should not occur on that copier model, or the solution given in the manual was incomplete. To overcome the limitations of the service manuals, technicians would gather over coffee and exchange stories from the field. Gradually a parallel service manual, based on real world experience, emerged, printed not on paper, but in the collective knowledge of the technicians themselves.

Although their members usually "*work on their own individual projects or teams and have few recognisable shared goals*" [Millen02], CoPs foster a strong sense of belonging, of being 'in on' something. Establishment of trust between members is paramount and, as a result, CoPs tend to be weakened if there is frequent membership 'churn' [Sharp97]. CoPs typically have a strong sense of place [Dourish01]; they form around some Schelling [Schelling60] or focus point. Even the Xerox technicians drank their coffee in 'their' corner of a warehouse. While many CoPs are now, primarily, virtual communities, with members 'meeting' in online newsgroups or web bulletin boards, the most successful remain those whose membership hold regular face-to-face meetings [Sharp97]. The role of community ritual or etiquette is also emphasised – Orr's Xerox workers, we are reminded, swapped stories over cribbage games – and members perceived as being at the periphery of the group may be branded *lurkers* [MacDonald03], deriving benefit from the group's expertise without contributing to it.

Were this attitude to prevail, opportunity for cross-community interaction would obviously be severely restricted. Scarbrough [Scarbrough96] warns of communities of practice becoming "*inward-looking and unable to recognise the contribution and knowledge of other groups*"; whilst those "*…within disciplines can communicate and collaborate with colleagues globally, they may fail to collaborate across disciplines, even at a local level*" [Swan02]. It is at the intersections between CoPs, though, that radical, or 'discontinuous' [Veryzer98] innovation most often begins [Blackler95]; "*innovations are rarely conceived within the framework of a single community of practice*" [Miller99]. It is the lurkers, or rather the *boundary members*, that are most often the innovators. It is also likely that these boundary members will be the most effective in performing the knowledge-base interpretation role discussed above. Scarbrough concludes, therefore, that "*the development of a community-based approach to innovation needs to be able to exploit the advantages of communities… while overcoming their tendencies towards social closure*" [Scarbrough01].

If a CoP *is* successfully anchored into a wider organisational network, experience has shown that it becomes a resource for the entire employee base. An example of this within an innovation-led organisation can be found in the description given by Seely Brown and Solomon Gray [SeelyBrown95] of the phase lock loop (PLL) engineer community at National Semiconductor. Developing a word-of-mouth reputation for excellence, the community was increasingly called on by engineers throughout the company to assist with PLL-related design issues. In doing so, it increased its own body of knowledge through tackling a broader range of problems than it would otherwise have encountered. Furthermore, establishment of effective outward-reaching social ties enabled the CoP to develop a much clearer sense of common purpose internally. Initially an ad hoc group of twenty-or-so similarly interested individuals, the PLL CoP gained a formally recognised physical presence - a “PLL place”, created in lab space donated by a product group whose work it benefited and acquired funding in its own right to develop projects outwith the control of the product groups.

2.3.2 Communities of Interest

As Miller observed, while established Communities of Practice foster the expertise needed to progress an innovation, they are not, typically, environments in which radical innovation takes place. Instead, Scarbrough tells us, “*innovation involves the creation of small, transient communities which are able to provide a context for sharing of knowledge, yet are equally able to interface with existing communities of practice*” [Scarbrough01]. Solomon similarly identifies coalition building as the essential first step in achieving innovation from personal ‘inspiration’ [Solomon98]. Fischer and Oswald term this type of community a Community of Interest – a “*community of representatives of communities*” – who have come together from different fields to jointly tackle a problem or challenge [Fischer01]. That is not to say that the community necessarily results from an urgent need or task; the initial motivation for forming a CoI may, as Cornejo says, simply be “*curiosity*” or “*debate*” [Cornejo03], inspired more by collective personal interest than professional obligation. While CoPs are characterised by the collective expertise of their membership in a single field, CoI members are at once experts and novices; communicating their knowledge to others in the community as experts, but also, as novices, gaining knowledge from other members who have expertise beyond their own particular field [Fischer03]. Fischer notes that this “*asymmetry of ignorance*” can be exploited as a creative stimulus, through which CoIs gain greater potential to innovate than a single CoP.

2.3.3 Intensional Networks

Linking these communities, the netWORK study conducted by Nardi et al. [Nardi02] identified a looser worker-network model, based more on social relationships than shared competence or activity. *Intensional* networks are more heterogeneous than communities of practice or interest – they include people who have little in common professionally, but who either know each other personally, or have been recommended by colleagues. An intensional network is a personal network. In other words, while many people belong to the same community of practice, and are externally identifiable as members of such, each individual maintains and nurtures their own intensional network. Nardi coined the phrase 'intensional network' in recognition of the stresses individuals face in acquiring and renewing contacts and in tracking their activities and positions.

Intensional networks are founded on the 'six degrees' phenomenon; the oft-cited rule of thumb that states that any two people in the world are connected by a chain of, at most, five other people. Though it has long been commented that we live in a 'small world', the first significant scientific study of this effect dates to the 1960s. Travers and Milgram [Travers69] conducted a simple experiment:

> "An arbitrary 'target person' and a group of 'starting persons' were selected, and an attempt was made to generate an acquaintance chain from each starter to the target. Each starter was provided with a document and asked to begin moving it by mail toward the target…
>
> It was stipulated that the document could be sent only to a first-name acquaintance of the sender. The sender was urged to choose the recipient in such a way as to advance the progress of the document toward the target; several items of information about the target were provided to guide each new sender in his choice of recipient."

For those documents that stated the target's occupation, the mean number of people through whom the document passed before reaching the target was 4.6; *six degrees of separation*.

Common sense tells us that Travers' and Milgram's result is as deceptive in its simplicity as it is elegant. Not everyone is equal in importance in an intensional network, nor can every link be equal in strength. Granovetter [Granovetter73] noted the existence of "*strong*" and "*weak*" ties, defining the strength of a tie as "*a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie*"[5]. Interestingly, having a large number of weak ties in a network was seen to be a catalyst for innovation and creativity.

Though this tallies with Ferdinand Toennies' dichotomy of strongly tied agrarian *Gemeinshaft* (community) and loosely tied industrial *Gesellschaft* (society), it is not immediately obvious why scarcely knowing a large number of people is a creative or inspirational situation. Nardi's ethnographic study offers an explanation. She found that weak ties could quickly be strengthened when that part of an individual's activity was centred on that part of their 'sub network', while other ties outside this centre of activity were allowed to weaken to the extent that effort was required simply to remember them. Persistent strong ties, on the other hand, demand constant attention. Networks with large numbers of weak ties had the most potential for rapid reconfiguration and so were the most flexible.

Many of the individuals involved in Nardi's study worked in the creative industries and it is, perhaps, unsurprising that social or intensional networks play an important role in an area such as this; one that is dominated by a mix of self-employed freelancers and small, collegial enterprises – that is, consensus-driven

[5] Friedkin [Freidkin82] offers a simplified definition, whereby the strength of a tie is a measure of information flow between the tied nodes.

and staffed by individuals who are "*equals in their levels of expertise but… specialized by area of expertise*" [Waters89]. It is argued, though, that the personal networks described by Nardi are also powerful enablers of innovation in ostensibly rigidly hierarchical organisations, where their informal, status-agnostic ties facilitate cooperation across management-imposed hierarchical boundaries [Conway98] [Kautz97] [Lazega01]. As Majchrzak et al. [Majchrzak01] put it, "*it's hard to cold call someone to find a new solution*"; awareness of connecting networks provides both a social reason for an expert to reply to a request for help that they might otherwise ignore (for example, the knowledge that they have a mutual collaborator) and a criteria by which an individual seeking assistance can assess the trustworthiness of an expert. The importance is also acknowledged of informal contacts and personal relationships in "*breathing life*" into formal officially-sanctioned collaborations [Conway98].

2.4 Social Ontologies

If software support is to be provided for social networks and Communities of Interest, the relationships and classes that they involve must be modelled. Efforts by KM traditionalists to capture, manipulate and share knowledge electronically have motivated development of a means by which the conceptual elements of this knowledge can be represented and interrelated [Benjamins98] [Uschold96]. An ontology is a "*specification of a conceptualization*" [Gruber93]. It provides a shared and common understanding of a domain, by describing concepts and their relationships to each other [Genesereth87] [Zdrahal01]; opinions of just what a domain conceptualisation must include to constitute an ontology do vary, but Deborah McGuinness suggests that, as a minimum, an ontology must provide a "*finite controlled vocabulary*", "*unambiguous interpretation of classes and term relationships*" and "*strict hierarchical subclass relationships between classes*" [McGuinness01]. While ontologies are still most strongly associated with traditional IT-based KM applications, a number of ontologies have been developed that encapsulate aspects of the social and collaborative relationships within organisations.

One of the most widely recognised social ontologies is also one of the simplest. FOAF, or Friend Of A Friend [Dumbill02], is an open, community-driven RDF (Resource Description Framework) vocabulary for machine-readable representation of individual's personal attributes – their name, organisation, homepage URL and so on – within web homepages. As its name suggests, a key aspect of FOAF is its ability to indicate personal networks. Its 'knows' relationship allows an individual to reference others with whom they are familiar; these relationships can then be mined by FOAF browsers to produce personal network maps. A proposed extension to the core FOAF specification provides additional relationships such as 'worksWith' and 'participantIn'. FOAF's ability to represent interests and community involvements is, however, currently quite limited; an indirect means to declare current personal interests is provided, allowing individuals to reference documents tagged with FOAF 'Topics', and a project or group can have associated with it an overall 'Theme', but description of an individual's specific involvement is limited to a distinction between 'PastProject' and 'CurrentProject'.

Social ontology components can also be found within broader ontology specifications. SHOE (Simple HTML Ontology Extensions) [Heflin99], the Semantic Web Portal Ontology [Moller04] and the AKT Reference Ontology[6] each include both person-oriented concepts and representations of documents, events and administrative structures found within an organisation. The Semantic Web Portal Ontology, notably, combines FOAF's representations of people and organisations with additional original components. However, while reuse of existing ontology where possible is a laudable aim, the Semantic Web Portal Ontology's authors acknowledge that the fluid nature of FOAF demands constant realignment of their own vocabulary. The AKT Reference Ontology, too, has

[6] www.aktors.org/ontology/

evolved continually as it is subject to ongoing academic discussion. While optional extensions at the periphery of an ontology are to be expected, radical restructuring of key concepts would be difficult to accommodate were the ontology to be implemented within a practical application.

Although an ontology should *represent* the structures, relationships and classifications that exist within an organisation, it must not constrain them. In the AKT Reference Ontology, for example, two areas are evident where additional flexibility would be required were the ontology to form the basis of an innovation support platform. In it, occupational roles are specified as subclasses of its 'Person' class, effectively limiting an individual to a single role chosen from those specified by the ontology's authors. Topics from which individuals' interests are compiled are also hard-coded into the ontology; by contrast, not only would an ontology underpinning an innovation support platform be expected to serve unanticipated future uses of the platform, but use of the platform would itself be expected to generate new topics, as the innovation that it supported came to fruition. Each of these broader ontologies also shares FOAF's limitation in respect to representation of community involvements; the Ontocopi Community of Practice visualisation application [Alani03], discussed in Chapter 3, highlights the inadequacies of the AKT Reference Ontology in this respect.

The ontologies discussed here are concerned, primarily, with public representations of a person or organisation and designed to provide generic, reusable, domain taxonomies to which multiple separate and potentially diverse applications can subscribe so that they can freely exchange information. This ambition necessitates domain conceptualisations that can 'safely' and usefully be exploited and viewed beyond their host organisation by unknown parties. Our application, though, will demand a stable representation for internal use that is both transparent and flexible. In Chapter 4, these issues are revisited as an ontology is constructed specifically as the basis of an innovation support platform.

2.5 Summary

Innovation and knowledge are intrinsically linked, but while IT-based Knowledge Management strategies have been deployed successfully in a range of organisations, but they have not been found to support innovation. Although it is often said that this is because innovation relies on tacit knowledge that is not amenable to knowledge management, a more significant factor is the role in innovation processes of face-to-face negotiation of common ground and exchange of situation-specific implicit knowledge. Recognising this, a community-based approach to knowledge management has been proposed to support innovation, which emphasises support for the communities of practice, communities of interest and intensional networks through which face-to-face collaborations take place. Ontologies have been introduced as a tool rooted in traditional KM that can also support this community-based approach. In the next chapter, we will examine a range of recent community-aware workplace technologies and consider their suitability for innovation-led organisations.

Chapter 3

Community-based KM Tools

While dissemination of established procedures and technologies can be achieved with an IT-based knowledge management approach, innovation relies heavily on the face-to-face transfer of knowledge between individuals, through a combination of established Communities of Practice, task-oriented Communities of Interest and personal intensional networks. In this chapter, we will survey a range of technologies that have been developed to nurture or exploit these networks and consider their suitability for an innovation-led organisation.

3.1 Network and Community Visualisation

Once, if you wanted to understand the relationships between employees in an organisation, you would visit its personnel department and ask to see their organisation chart. This chart showed who worked with whom, in what team, under which manager (organisational charts are nothing if not hierarchical). More than that, it was a tool for planning, for control [Krebs02]. Relationships were made and unmade not by the employees themselves, but by a personnel manager drawing and rubbing out lines of ink, sticking on labels and peeling them off again. In modern innovation-led organisations, though, the real organisation chart is redrawn constantly by the evolving, employee-managed networks and communities through which innovation takes place.

Social network mapping has its origins in the hand-drawn sociograms produced by Jacob Moreno in the 1930s. Computer-generated visualizations for academic use followed in the 1970s [Freeman00] and, more recently, these have become widely used within commercial and governmental organisations, both as general-purpose management tools and as a component in the technical armoury of criminal and military investigative analysis bureaux[1]. We will examine four interpersonal network applications that are representative of recent research in this field that is relevant to innovation-led organisations: ContactMap, Social Circles, Referral Web and ONTOCOPI.

3.1.1 ContactMap

ContactMap [Nardi01], shown in Figure 3.1, is a network visualisation and contact management application, developed by Bonnie Nardi and colleagues when at AT&T Labs. In operation, the main ContactMap window resembles a desk top onto which pages from the user's address book have been scattered, then arranged according to how recently and often they have communicated. Stronger contacts or contact clusters are placed towards the centre of the desk top, while weak ones are arranged around the edge. The user bootstraps ContactMap by running an

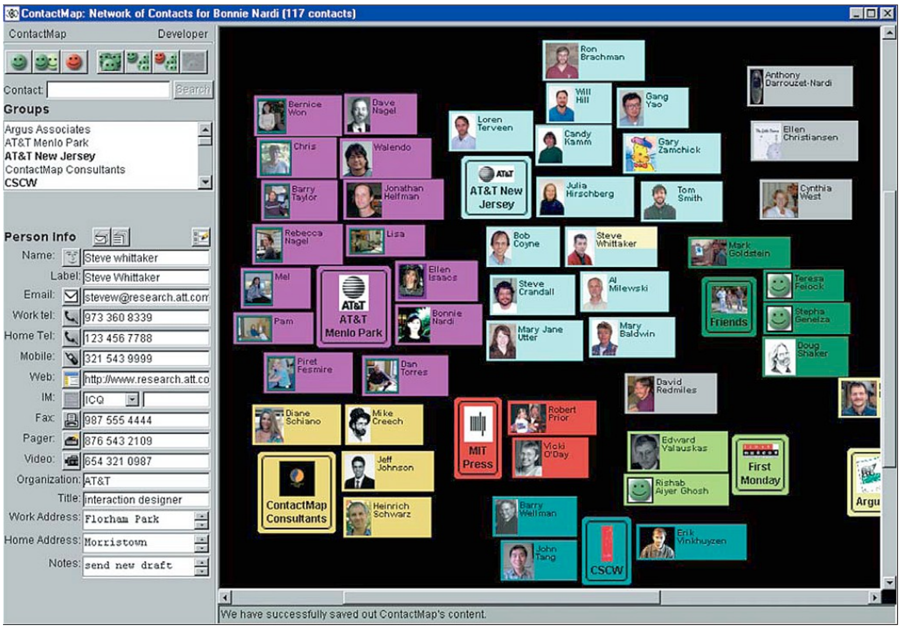


Figure 3.1 – ContactMap

---

[1] www.i2.co.uk

email analysis tool, which creates a model of their immediate social network by ranking each contact in their email archive according to recency and number of communications. They are then required to manually identify those contacts that they wish to include in the visualisation and cluster them, where appropriate, into project, organisational or family groupings. Contacts can be manually annotated with a photograph of the individual and additional contact details such as phone or fax number; short text notes can also be attached, giving a limited task management function.

ContactMap offers a straightforward means for the user to manage their immediate social circle, but as tool for community-based KM, it has a number of obvious limitations. The ContactMap model is not automatically updated, so to add a new contact the user must either manually create a new node, or re-run the email analysis tool. As the application does not represent a contact's current activity or interests, it cannot bring to the user's attention individuals who have become more relevant to their current work. Furthermore, as the contact model only includes individuals with whom the user has had direct email communication, ContactMap does not present new opportunities for collaboration or allow colleagues' social circles to be exploited.

3.1.2 Social Circles

Whereas ContactMap bases its social model on email communications directed to a single individual, Marcos Weskamp's Social Circles [Weskamp04] "*intends to partially reveal the social networks that emerge in mailing lists*". Synchronizing directly with the mail server every 5 minutes, the Social Circles Flash client (Figure 3.2) presents a near-live representation of the social dynamics within the list. Like ContactMap, Social Circles uses a spatial metaphor to depict the social proximity of an individual. Here, however, it is the individual's importance within the group, rather than to a single acquaintance, that is shown.

Each node depicted in the visualisation represents a participant in the list. Whenever the participant sends a message, their node increases in diameter. If their message receives a reply, their node is moved towards the centre of the visualisation, mirroring their increasing centrality to the group and a line is drawn linking their node to that of the respondent, to signify the line of communication.

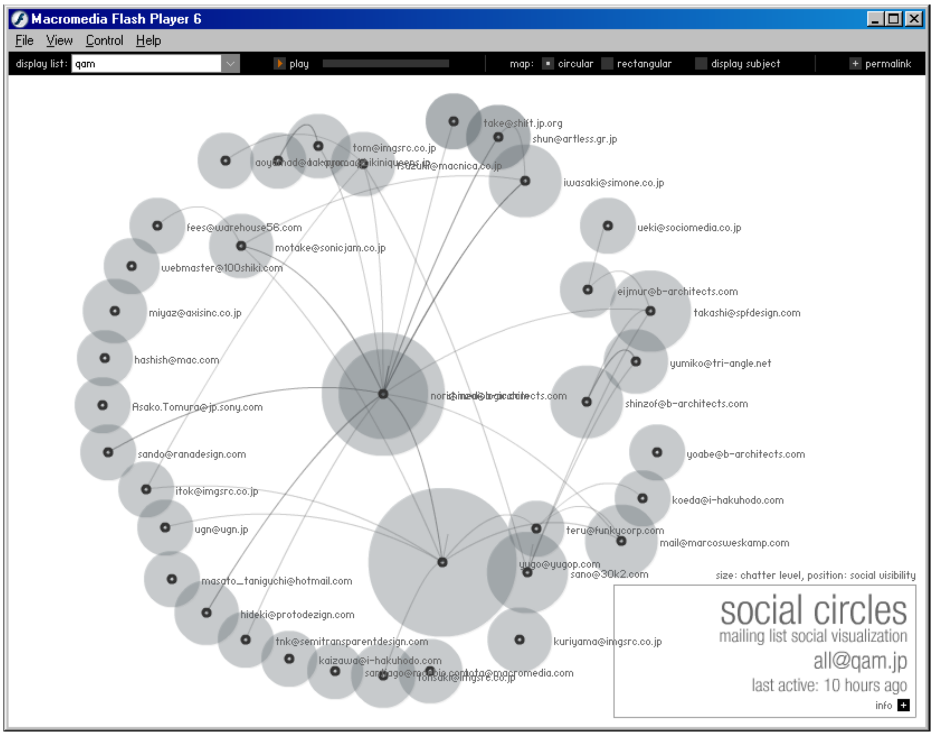


Figure 3.2 – Social Circles

Social Circles demonstrates that email logs can be used not only to determine a single individual's immediate social circle, but also to observe the more complex power politics and sub-groupings that emerge within a community. It also highlights two drawbacks of this approach: the application can only visualise communities of which the user is already aware and cannot represent the interests of those non-participating boundary members that are such important stimulators of innovation.

3.1.3 Referral Web

Referral Web [Kautz97] was developed by Henry Kautz and Bart Selman, also at AT&T Labs. It is a tool for unconstrained social network exploration that "*uses the co-occurrence of names in close proximity in any documents publicly available on the WWW as evidence of a direct relationship*". Using this document co-reference model, Referral Web can identify the social circle surrounding a single named individual, or determine the strongest social path between two named individuals. In each case, the documents linking pairs of individuals can be listed. Figure 3.3 illustrates Referral Web's visualisation of the social chain linking Nigel Shadbolt to Marvin Minsky, with the collaboration between Shadbolt and intermediate node Austin Tate detailed. Referral Web also has an expert finder function that ranks individuals according to the number of publications they are linked to that contain the specified keyword; the search can be user-limited to a specified number of degrees of separation from a named individual.

Whenever the system encounters a previously unknown individual, a web spider is launched to determine their social network to 3-degrees-of-separation. The spider uses the Altavista search engine to retrieve web documents that contain the target's name. Natural Language techniques are then used to extract any other individuals' names that appear in the corpus. This list of possible contacts is refined by thresholding the Jaccard coefficients between the target and each of the other names; the Jaccard coefficient simply being the number of papers retrieved that name both the target and the potential contact, divided by the number retrieved that reference *either* the target or potential contact. The spidering process can take up to 24 hours, depending on how many of the target's social contacts are also new to the system – as a result the demonstrator application presented is restricted to searches within a pre-introduced 10,000-strong list of Artificial Intelligence researchers. Clearly, though, if the scope of the system was limited to a single organisation, both the user-base and document corpus involved would be significantly smaller and this overhead would be drastically reduced.

As its social model is extracted from web-published documents, Referral Web does not need to be bootstrapped by users. Individuals are visible to the system by virtue of their documentary presence online, even if they have never used it. For a system that aims to provide an unbounded representation of global expertise and established social contacts, this is essential; less so, though, in an intra-organisational context, where organisation-wide subscription to a single software tool can be mandated. Of more concern in this context is the significant latency that results from a dependence on post-hoc declarations of collaboration in formal publications – a collaboration between individuals that have not previously published together might only become apparent months after it has ended. This dependence imposes further limitations: the inferred relationship model obtained by Referral Web cannot, for example, reliably distinguish between a series of collaborations over a period of years and multiple references to a single joint-presentation, or identify long-term but undocumented informal collaborations.

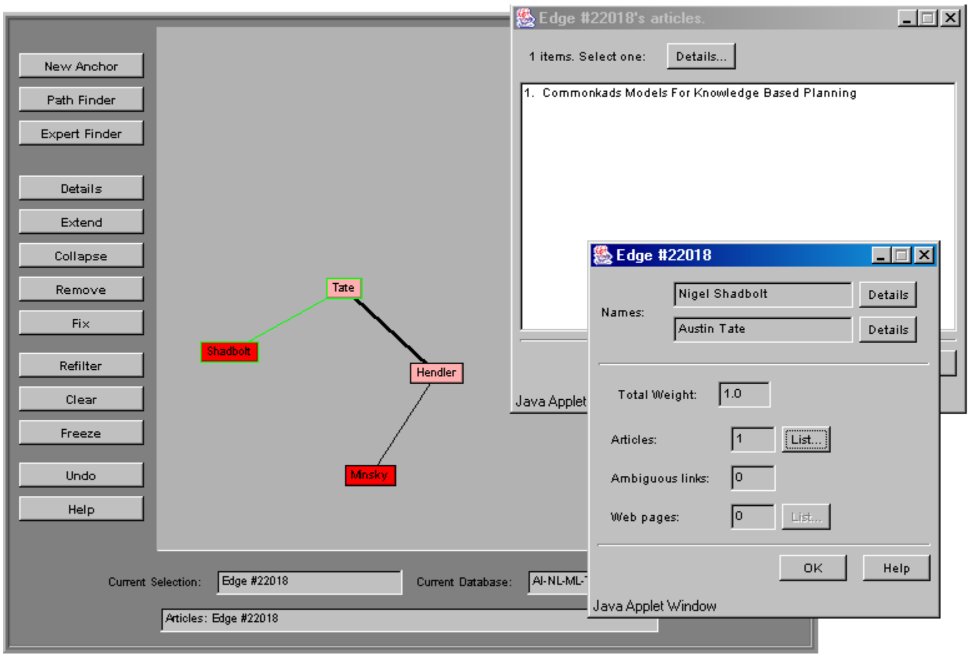


Figure 3.3 – Referral Web

3.1.4 ONTOCOPI

ONTOCOPI (Figure 3.4) [Alani03] is a Protégé [Eriksson99] application intended to identify and visualise Communities of Practice. Unlike the other applications surveyed, ONTOCOPI does not seek to infer relationships from email or web evidence. Instead, it makes use of a pre-populated domain conceptualisation, or ontology, in which such relationships are explicitly represented. In contrast to the other visualisation tools discussed, therefore, ONTOCOPI has access not just to the fact of collaboration but also aspects of the scope. Concepts modelled by the ontology include document authorship, event attendance and project and organisational membership.

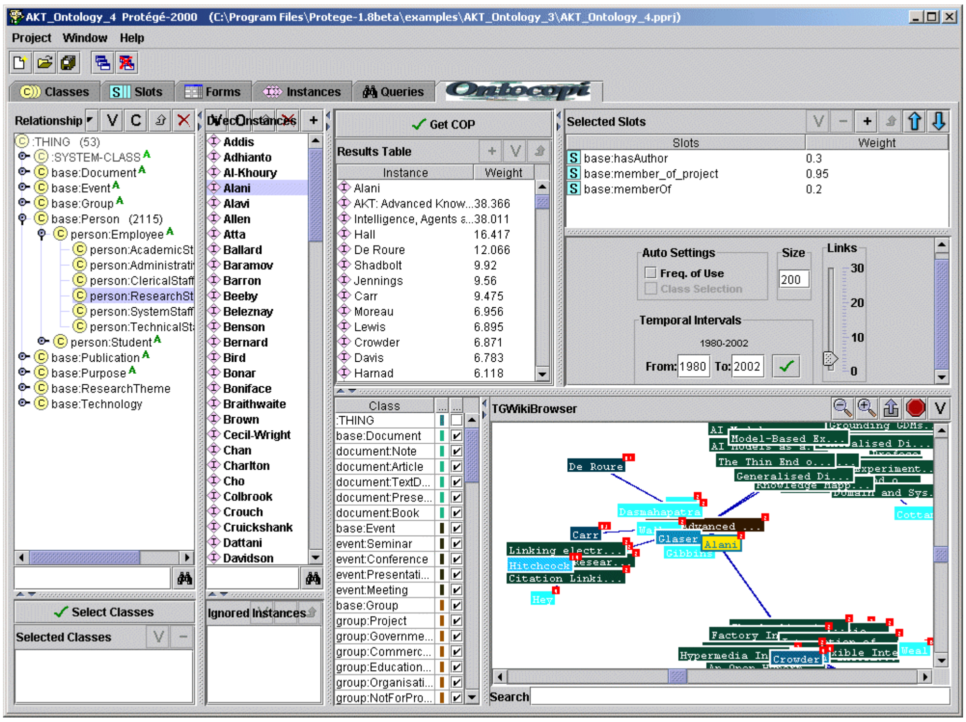


Figure 3.4 – ONTOCOPI

The user can select which of the relationships contained in the ontology they wish to be used to construct the community model; additionally the influence of each relationship type can be varied by adjusting a weighting factor. Alternatively ONTOCOPI can select and weight the relationships to be used automatically, based on the incidence of each relationship type in the ontology.

ONTOCOPI then applies a breadth first, spreading activation search, traversing each of the relationships and scoring their strength, until the link threshold is reached. The resulting interpersonal chains are visualised using a hyperbolic tree graph and the end points are also ranked and presented in tabular form. As the ontology includes a temporal component, ONTOCOPI can also show how an individual's collaboration network has evolved over time; a slider control allows the user to limit the temporal scope of the collaboration model.

The relationships that ONTOCOPI uncovers are perhaps more accurately characterised as *collaboration networks* than Communities of Practice; in practice ONTOCOPI is unlikely to reliably identify CoPs because the ontology used, the AKT Reference Ontology, has only limited capability for modelling of individuals' interests or expertise and, as Millen [Millen01] indicated, the professional collaborations that the ontology does capture rarely occur between members of a single CoP. Though ONTOCOPI's developers concede the *problem of brokers or boundary objects* [Alani02], their suggestion that this be addressed by *filtering out noise* is inappropriate in our application where boundary individuals are an important asset. Furthermore, while ONTOCOPI regards discontinuation of document- or project-level collaboration between individuals as evidence that they are no longer members of the same CoP, this might instead represent the maturation of a field, where internal development and evolution have subsided and members' attentions shifted to engagement with others outwith the CoP.

Despite these shortcomings, ONTOCOPI does demonstrate that ontologies can be used to derive richer organisational network models than can be obtained by passively monitoring web documents or email logs while maintaining separation between the application and the organisational conceptualisation on which it is based.

3.2 Information Delivery Systems

In the previous chapter the importance of communication within and between groups and communities was emphasised. Direct, person-to-person communication is well supported [Koch03] by telephone, email and, most recently instant messaging (IM) [Nardi00] technologies. Such technologies, though, are less well suited to speculative or loosely targeted communication. While email remains for communities the primary means of information exchange, the receipt of large quantities of potentially irrelevant email is seen as a significant burden on employees. Noting that email has been described as the ‘killer application’ of the Internet, Ducheneaut and Bellotti suggest that, in reality, “*email is a serial-killer! It is seriously overloaded and used to manage a variety of tasks it was not meant to support seriously*” [Ducheneaut00]; Whittaker and Sidner [Whittaker96] have made similar observations. Goecks and Cosley [Goecks02] also found that individuals are often reluctant to send speculative messages by email as it is perceived as adding inappropriate “*extra weight or importance*” to the communication. Recognising the limitations of email, researchers have built a number of alternative systems specifically for the exchange of these speculative or loosely targeted messages.

3.2.1 NuggetMine

One of the simplest community information technologies developed is NuggetMine, from University of Minnesota [Goecks02]. NuggetMine provides a forum for the exchange within a community of small ‘nuggets’ of information, such as URLs or snippets of news; items which are considered too trivial for email or whose target audience is not clearly defined, but would otherwise be lost to the community.

To submit a nugget, the user simply drags text or a URL to the lightweight desktop client; they may optionally specify dates between which the nugget should be displayed. The nugget is submitted to a central server, where it is titled automatically, using either the first few words of submitted text, or, in the case of a URL, the title of the referenced web page. The server characterises the nugget by running the Term Frequency inverse Document Frequency (TFiDF) algorithm on the nugget text or referenced web page text to generate a list of keywords. If the user has submitted the nugget in reply to one displayed on their client (see Figure 3.5) this relationship too is recorded.

The desktop client includes a nugget headline window, which receives from the server a stream of titles of the most recently submitted nuggets. Clicking the currently shown nugget title opens its content in a browser window. Additionally, the user is presented with a short list of related nugget titles, derived by the server by matching the TFiDF characterisations.

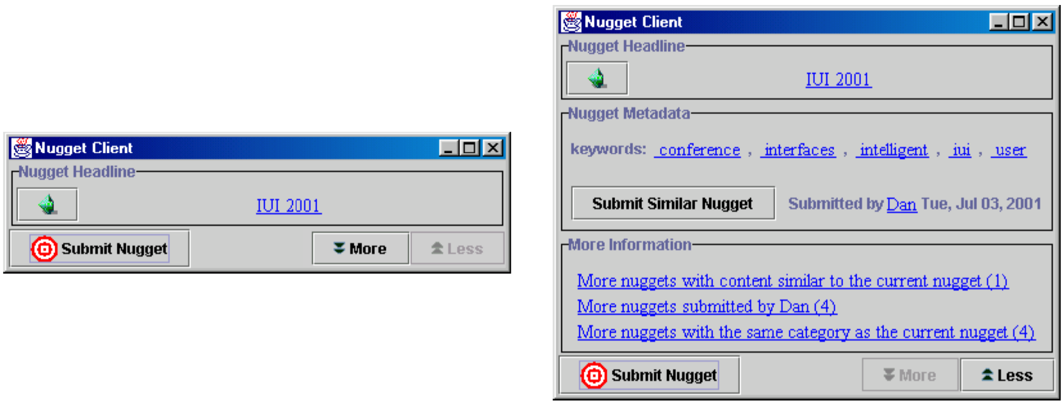


Figure 3.5 – NuggetMine Client

NuggetMine indicates that a profile-based approach can be used to deliver information, but there are obvious drawbacks in demanding that users manually identify relevant nugget clusters. The reliance on machine-generated titles catching users' attention as they scroll past means there can be no guarantee that information will reach the most appropriate individuals; if it is important that a message be conveyed to a particular individual or group of individuals, a duplicate must be sent using a conventional absolute-addressed messaging technology.

3.2.2 GroupCast

Accenture's GroupCast [McCarthy01] is a public information display that presents material relevant to passers-by. The system is mainly used to present general interest information snippets such as horoscopes, stock reports and news headlines, but it also allows users to post group events flyers, general community announcements and NuggetMine-like suggestions of 'what's cool'; these 'what's cool' items can optionally be addressed for the attention of specific named users.

Each user create a personal profile that indicates their interest in different classes of notification; subscription is on a per-channel basis, with the exception of the event notifications where an individual can select which events they wish to be presented with information about. Accenture Labs employees wear an electronic ID badge that is tracked via a ceiling-mounted array of infra red sensors, allowing their position within the building to be pinpointed. GroupCast monitors the proximity of individuals to the public displays and, by referencing these individuals' personal profiles, selects material for presentation that is most likely to be of interest to those nearby. Thus, GroupCast extends the profile-based approach of NuggetMine by, albeit simplistically, matching information and user profiles.

3.2.3 myPlanet

myPlanet [Kalfoglou01] is an ontology-driven personalised news service developed by the Open University's Knowledge Media Institute. It provides an archive repository for 'press release'-style news items concerning academic events, which users can search according to selected key concepts drawn from an underlying organisational conceptualisation.

News items are submitted to the system's OntoPlanet [Domingue99] server as plain-text email messages. An automated Information Extraction (IE) component parses the content of the message, using pre-defined templates to identify salient concepts such as event attendance or location. These concepts are used to pre-populate an annotation form, which a "*knowledge engineer*" or news editor completes to fully describe the item. Readers receive a selection of stories that the system identifies as relevant to concepts in which they have expressed an interest. These concepts can be saved as a user profile, allowing a customised web news bulletin to be provided. As the only input an author makes to the system is the text of their news item (and, optionally, an attached image), they do not themselves have direct control over the eventual visibility of their item, so project- or group-internal news items are not supported.

The approach that myPlanet adopts is appropriate for formal public news releases, but is less suited to impromptu exchange of the lightweight snippets and pointers that NuggetMine supports; this is reflected by the relatively low reported proportion of OntoPlanet users that are contributing authors (13 out of 480) [Domingue99].

3.2.4 Drehscheibe

Munich Technische Universität's Informations Drehscheibe (Information Turntable) [Koch03] application was developed to support communication in an academic environment. The central concept of Drehscheibe is a user-managed community hierarchy, maintained in a mySQL database. Using the system's personalised web browser interface (Figure 3.6), any user can create a community and position it in the existing hierarchy. Communities can be declared 'open' to anyone or 'closed' – in which case membership requests must be approved manually by a community administrator. In addition, there are specific privacy settings for the community information board, allowing both the publishing and reading of material on the board to be restricted to community members.

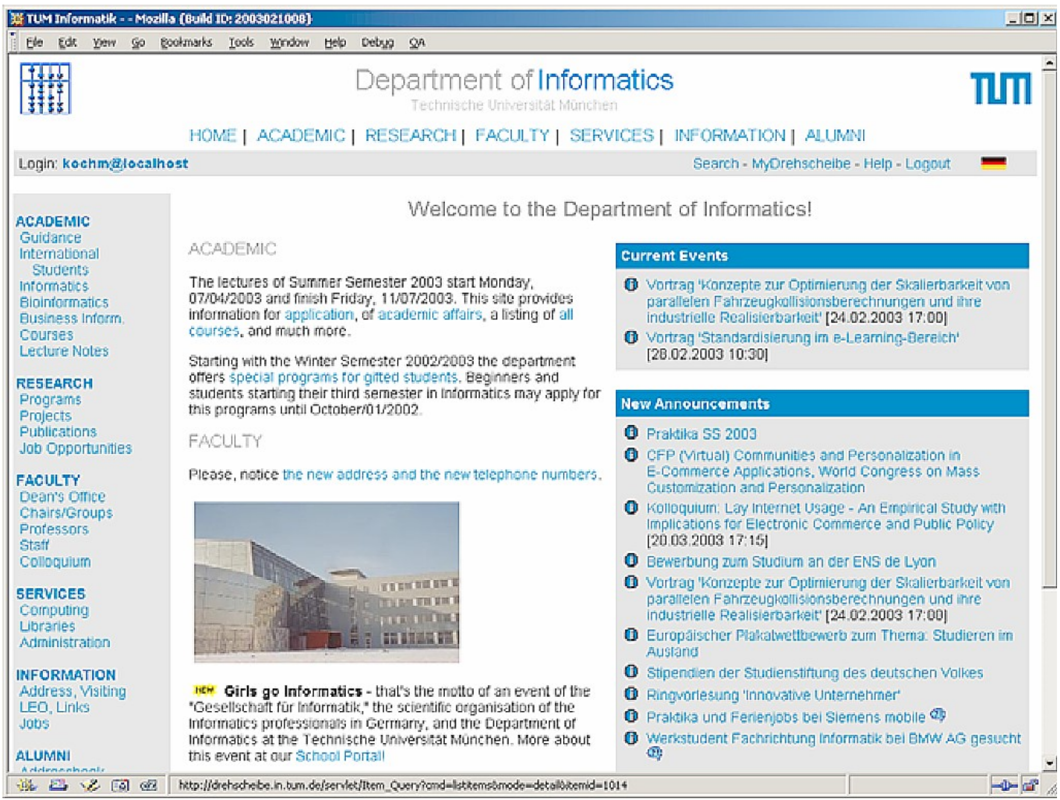


Figure 3.6 – Drehscheibe

Day-to-day use of the system is also via the web browser interface. Messages can be posted to one or more communities; a range of message templates are provided to allow structured submission of events, web URLs, project announcements and so on. Again, start and end times for the display of the message can be specified. A description and membership list of each community can be obtained; individual users' profiles identifying their community memberships and hand-rolled 'buddy lists' can also be retrieved.

While Drehscheibe introduces a richer community structure, Koch reports that users were less keen than had been anticipated to create new communities. It is clear, though, that Drehscheibe does not easily allow potential members for a new community to be identified, nor the existence of a new community to be advertised to those who might be interested without 'spamming' many others. There is no means for individuals to declare interests other than by joining a community and individual involvements within a community cannot be described.

Furthermore, because the nature of the inter-community relationships is not explicitly stated, this structure cannot be exploited by the system to enhance community awareness; it is left to the user to judge, for example, whether a parent-child community tie represents a funding relationship or common intellectual ground.

3.2.5 Knowledge Pump

Like Drehscheibe, Xerox's Knowledge Pump [Glance98] too provides a hierarchical community structure, used specifically to exchange web site recommendations that are submitted using a lightweight Java client application and retrieved via a custom browser interface. Significantly, however, it also introduces the concept of social network awareness. Using the Java client, the user can nominate a list of 'friends' from the application user base. The system uses this list in selecting which site recommendations to present to the user; in addition to scoring up pointers from individuals with similar interests, it also weights more highly those that were submitted by the user's nominated friends.

3.3 Inter-Personal Awareness

Communities and networks form when individuals and communities come together. Individual and community visibility, or *awareness* is therefore essential in innovation-led organisations. Awareness is clearly a somewhat abstract concept to support, but for the purposes of developing electronic means to promote awareness within the workplace, a number of clarifications have been proffered. Greenberg [Greenberg96] regards awareness as "*the general sense of who is around and what others are up to*". Dourish and Bly [Dourish92] suggest awareness "*provides a view of one another in the daily work environment*", it is about "*knowing what activities are occurring, who is talking with whom*". For Gutwin and Greenberg (again) [Gutwin96], awareness is also an understanding of peoples' "*intentions*"; what *is going to* occur as well as what *is* occurring.

Stenmark emphasises that awareness is a two-way process, that "*individuals benefit both by being able to find knowledgeable colleagues and by being themselves identified as knowledgeable*" [Stenmark02]; what Dourish [Dourish02] identifies as 'mutuality'.

3.3.1 Inter-Personal Awareness Devices

A prominent aspect of organisational awareness research has been the development of Inter-Personal Awareness Devices, or IPADs; electronic devices that encourage awareness between physically co-located individuals. In 1988, Mark Weiser, then head of the Xerox PARC Computer Science Laboratory, first articulated his vision of a '3rd wave' of computer technology: *ubiquitous computing* [Wieser91]. While previous generations of computers have required the user to adapt to the computer's environment, ubiquitous computing resides in the human world. It is situated within everyday objects, embedded in the fabric of the built environment - and in the fabric of our clothing. Ubiquitous computing has enabled the development of mobile CSCW applications, of which IPADs form a significant focus.

Practical IPAD research covers a predictably broad class of application, encompassing dumb 'electronic ice-breakers' and more intelligent, contextually informed, profile-based collaboration support tools. There is concern, though, that, ultimately, any approach that relies on physical proximity will be unsatisfactory: "*when knowledge transfer is limited to random person-to-person encounters, this creates a likelihood that individuals will only reuse knowledge from those with whom the reuser shares physical proximity, or from knowledge contributors who make themselves readily available to others*" [Mazchrzak01]. Nonetheless, we will explore a representative sampling of IPAD research.

3.3.2 Simple IPADs

Borovoy developed the MemeTag [Borovoy98] to enhance interpersonal awareness within arranged social gatherings. MemeTags are computationally augmented nametags, worn around the neck. Initially each tag is loaded with a different *meme* – a short text phrase expressing an idea or opinion. When two people meet, each views the meme offered by the other's tag, and decides whether to copy this meme to their tag in place of the one they are currently displaying. Alternatively a tag-wearer can choose to introduce a new meme, by reprogramming their tag at one of the kiosks provided. As people circulate through the group, the number of memes in circulation is steadily reduced through a process resembling Darwinian Selection, leaving only those with most support in the group. This agreed view, which reflects the *community belief*, is displayed in real-time on a graphical *community mirror*. The tag and an example community mirror are shown in Figure 3.7. This work has now been commercialised as the nTAG[2], an icebreaker and messaging badge for use at conventions and trade fairs.

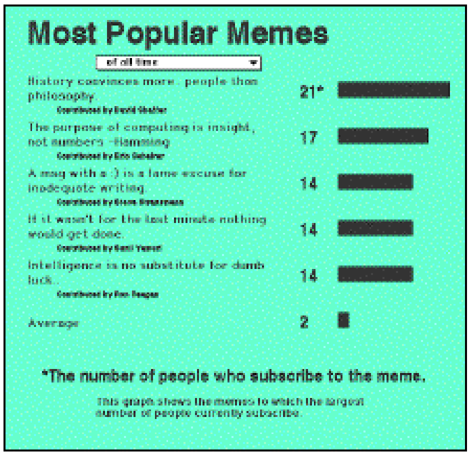


Figure 3.7 – MemeTag and Community Mirror

[2] www.ntag.com

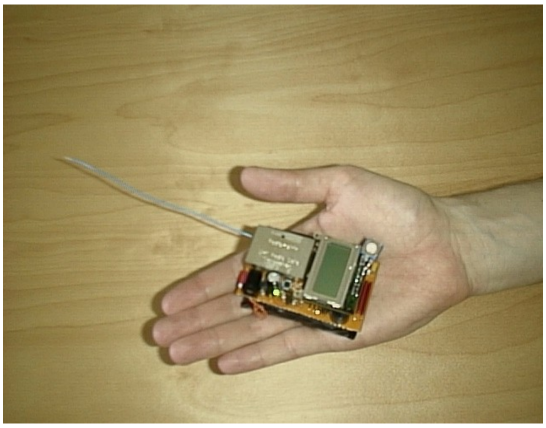


Figure 3.8 – Hummingbird

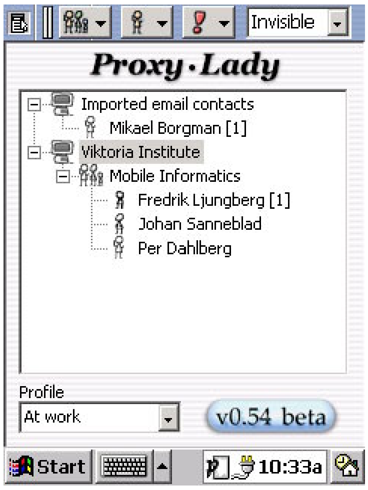


Figure 3.9 – ProxyLady

Hummingbird [Holmquist98] and ProxyLady [Dahlberg99] are non-infrastructured tools designed for ad hoc enhancement of awareness between people who already know each other. Hummingbird resembles a mobile *ICQ*[3] client, implemented using custom hardware (Figure 3.8). Each Hummingbird is assigned a *group*, and an *identifier* unique within that group. It announces its presence periodically, via a low power radio transmitter, and listens for other Hummingbirds within range. When a Hummingbird in the same group is detected, it emits an audio alert (i.e. it *hums*), and displays the identity of the other device on an LCD panel. The signal strength of the contact signal is also shown, giving an approximate indication of how far away the other device is. Although the Hummingbird was designed as a meeting facilitation device, in trials it was also found that it fostered a sense of togetherness or community, even when parties chose not to act on the alerts and meet physically, confirming that there is value simply in reminding individuals of the communities in which they are situated. Similar functionality is brought to the desktop by the BuddySpace IM client [Vogiazou03], in which traditional IM virtual presence indicators are superimposed onto geographical maps or office floor plans to provide at-a-glance awareness of the online status of individuals or project groups.

[3] www.icq.com

Dahlberg develops the role of meetings as opportunities for information exchange. ProxyLady is intended to support opportunistic *communication* – that is, unscheduled contact where one or both of the parties has anticipated the meeting and has prepared in advance an agenda of items to discuss. The current ProxyLady implementation supports discussion of emails. The system has mobile and desktop components. The desktop application, which connects to an IMAP mail server, retrieves the user's inbox and lists their current messages. The user may then associate individual emails with other "*candidates for interaction*" – other ProxyLady users. These associated emails and candidate identifiers are then transferred to a PDA running the ProxyLady mobile software, shown in Figure 3.9. When a candidate for interaction is encountered, the user is notified by an audible *proximity alert*, and the associated email is made available for them to refer to.

3.3.3 Proem

Proem [Kortuem01] is a JXTA-like platform designed specifically to enable interpersonal collaboration. It includes the "*notion of an individual as a person that owns and uses peers*". Proem includes a 'community manager' that tracks a peer's community membership, and operates as a gatekeeper to restrict access of non-member peers to resources held on the device. No one device has ownership of the community and any member of a community may confer membership on another device. Proem also supports 'shared dataspaces'. These are sets of data collectively owned and managed by a community and duplicated on each member of the community's device. As with Drehscheibe, a simple binary community membership model is used.

Though a Proem-based application could be used to disseminate information nuggets in a community of interest, the distributed way community membership is managed precludes the retrieval of any kind of membership list or community model – you are only made aware of a member of your community if they are physically co-located with you. While this would, in general, be a significant limitation in our application context, there may be circumstances where it would be appropriate to simulate this behaviour within a centrally managed system. It is conceivable, for example, that an expert may choose not to publicly advertise their expertise for fear of being inundated with emailed requests for help, but would be prepared to offer advice face-to-face to interested individuals they encounter.

3.3.4 hebb

hebb [Carter03] is designed to promote awareness between members of an organisation who share common interests. Interest profiles are generated by user-hosted 'interest sensors' that monitor the contents of the individual's received and sent email folders. On processing an email, the sensor first assigns the message a unique identifier, then eliminates commonly used generic phrases. A list of significant nouns and phrases is extracted using the Qtag[4] part-of-speech tagger; the list and message identifier are submitted to a central repository.

Like GroupCast, hebb combines user location-tracking with fixed public displays. The displays are paired with wireless access points; individuals are considered to be co-located with a display when their IPADs are bound with its corresponding access point. Whenever the set of users co-located with the display changes, their interest profile sets are retrieved from the repository and compared. A variant of the TFiDF algorithm is used to identify any significant pairings between different individuals' profiles.

Where significant inter-personal matches are found, the names of the individuals are shown on the public display, together with the matching keywords or phrases (Figure 3.10). As an aide-memoir or conversation piece, each user is

[4] www.bham.ac.uk/O.Mason/software/tagger

able to retrieve on their IPAD the message or messages that contained the matching terms. Non-present users with interests related to those common interests are also shown. To support impromptu communication if any of the co-located users wish to explore their mutual interest with other colleagues, there is an option to email all users identified as having the same interest; this is a one-shot function, the message is only received by those deemed relevant at the time the email was sent.

Again, a limitation of this application is its relatively simplistic conceptualisation of users' interests. hebb's developers report that, in practice, users have found that the common-interest topics identified by the system are often too 'common' to be useful. On the other hand, because the relationships between topics are not modelled, significantly *similar* interests may not be recognised.

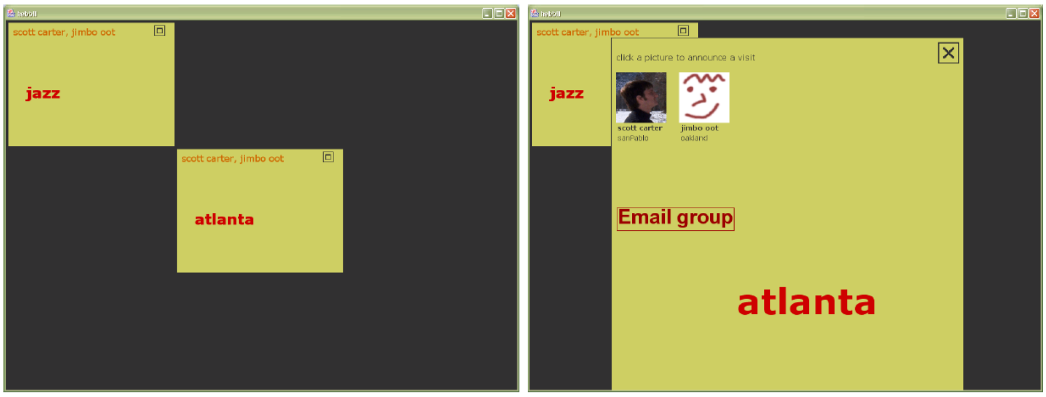


Figure 3.10 – hebb Public Display

3.4 A Piecemeal Approach

Although each component discussed addresses only one or two narrow aspects of community-based KM, no apparent thought has been given by their developers to how they might inter-operate to provide a more comprehensive tool set. That the lack of provision for communication and coordination between them will negatively impact the quality of service provided to the user, is, however, *quite* apparent. hebb and myPlanet, for example, hold estimations of users' interests -- hebb's drawn from their email correspondence, myPlanet's from explicit user declaration – but neither of these profiles can be accessed by NuggetMine; with no means to identify those snippets of information most likely to interest a particular user, it gives gold dust and garbage equal screen time. The informal communities maintained by Proem users are surely better indicators of Orr's self-organising clusters of expertise than chains of document authorship or formal project involvements, yet locked inside Proem's peer-to-peer architecture, they are invisible to ONTOCOPI and *its* users must rely on these more dubious approximations.

There are practical deployment and operational consequences too. Most of the tools use custom user profiles and organisational models, tying users into a particular set of applications and requiring that they maintain multiple representations of their contacts, activities and interests. With ONTOCOPI, the application and information model have been teased apart, but little is gained when this model cannot support the other components required. It is perhaps not surprising, therefore, that, of all the projects discussed, the only one that has achieved commercial success is intended for short-term use in the artificial, isolated, heavily managed environments of professional conventions and trade fairs.

In their attempt to liberate users from the intellectual and creative constraints imposed by IT-based KM solutions, these standalone, single-function community-based KM tools frustrate in other ways. Stenmark has emphasised the importance of a single cohesive organisational strategy to support awareness, information and communication: "*The communication perspective must not be isolated from the information and the awareness perspective*" [Stenmark02]. If this can be achieved with a coordinated software architecture, with a single, common, information structure, users should benefit from a consistent, seamless, very much more convenient solution that can provide substantially improved functionality while offering them the flexibility to assemble their own suite of client applications independent of the underlying platform.

3.5 Summary

A wide range of technologies has been developed that support aspects of community-based knowledge work, offering social and professional network visualisation, community-aware information delivery and person-to-person awareness. Unfortunately, the lack of coordination and interoperability and simplistic personal and organisational models limit their usefulness in supporting innovation. In the next chapter, a unified community-based KM platform will be specified that better addresses the unique requirements of innovation-led organisations.

Chapter 4

A Platform for Innovation

Although innovation is a knowledge intensive process, it does not respond well to traditional, centralised, information-based knowledge management. Instead, it is suggested that innovation is better encouraged by a community-based KM approach, which emphasises support for individuals' interpersonal networks. Though a range of tools is available that support some aspect of this community-based model, each has significant limitations. In this chapter, we will draw lessons from these existing tools in our consideration of the platform requirements suggested by the model of the innovation-led organisation presented in Chapter 2. We will then specify a flexible, holistic innovation-support KM architecture, Aware, based on a unified organisational conceptualisation, or ontology.

4.1 Awareness, Community and Communication

From the academic model of innovation-led organisations presented in Chapter 2, it is possible to identify three key areas that a KM architecture for such organisations must address: organisational *awareness*, *community* (of interest) formation and maintenance and information push, or *communication*. Awareness, at its simplest, is about "*how to bring people together… how to link people with others and with communities, that share similar interests*" [Hattori99]. The challenge awareness poses for a software tool is that "*people with similar interests often do not communicate about these interests because they are unaware of the relationship. As a result, many topics that people share an interest in remain unexplored*" [Carter03]. Peer-to-peer Inter-Personal Awareness Devices such as Proem [Korteum01] offer a partial solution, but can only introduce people that are physically proximate. If two individuals never meet, or only do so before they have picked up similar interests, they will not be made aware of each other.

Centrally maintained social network representations can enhance awareness of relationships with those beyond an individual's immediate social circle. They can be invoked step-by-step, as Nardi suggests, to enable indirect contact with an individual, or simply quoted to that person directly as a token of social introduction [Kautz97]. ReferralWeb can identify intensional networks to a named individual, but this requires that the user already know the identity of the specific person that they should contact. ONTOCOPI offers a predicted CoP, but still the user must be able to provide the identity of someone closely associated with the field. As the importance has been stressed of ensuring that CoPs are readily accessible by the wider organisation, a better solution would combine network visualisation with traditional 'expert finder' capability to allow social chains to any individual sharing a particular interest to be identified. Clark, Monk and Dahlbom have argued the importance of common ground and shared practice, but this is not addressed by any of the tools surveyed. A straightforward mechanism should be provided by which the user can be made aware of individuals that share common ground with themselves *and* have expertise or interest in another specified area.

Regarding community, two issues merit particular consideration: the ease with which new communities can be established and the extent to which existing communities are integrated into the wider organisation. Koch's experience with Drehscheibe suggests that users are reluctant to seed new communities electronically if they feel that doing so will publicly disrupt some established order or hierarchy. As communities of interest are often formed speculatively, or out of "*curiosity*" [Cornejo03], it is vital not only to ensure that the procedure involved is not too onerous, but also that the perceived commitment required is minimised. Nonetheless, even during its "*coalition building*" [Solomon98] phase, the community must be clearly visible to individuals with relevant interests. In the physical world, probably the lowest-effort, lowest-commitment way to communicate, raise awareness of and gauge interest in an idea is to pin simply-produced paper flyers to bulletin boards.

As Geyer and Cheng [Geyer02] point out, communities follow a natural progression from speculative communication, through initial open, fluid, team building and gradual refinement of focus, to formal project group with stable membership. This progression is neglected by the tools surveyed; none offer any means to classify the stage of development of a community. Once formalised, it is important that the community remain integrated with the wider organisation. Carter et al. found that "*members of groups tend to have strong connections to other members of the same group… but significantly weaker connections to other groups in their same organisation. This holds even when the other groups work on related issues or are related to the core group through an organisational construct*" [Carter03]. We have argued that '*innovation depends on the "integration of knowledge across disparate social communities"*'. Therefore, any system that exacerbates this situation, by encouraging internal boundary formation or Balkanization, or reinforcing tribal identity, will be counter-productive, yet in the tools surveyed community membership is characterised by a binary model – in other words, you are either in, or out.

In reality, however they might present themselves to outsiders, communities of interest are not single-minded monoliths and individuals' involvements with them are not defined by absolute subscription to a collective mission statement or umbrella title. Rather, they are focal points of individual interest and ambition; "*individuality of each member is preserved even in one community. That is, each member can have diverse objectives, even if all members share common interests in general*" [Hattori99]. A richer model of CoI involvement will reflect the "*individuality of each member*" *and* the shared "*common interests*", exposing both to the wider organisation.

As a community is formalised, the balance of information flow shifts. It evolves from a body predominantly concerned with compiling information from external sources to one that more often selectively communicates internally generated information. To increase awareness of community activities, the system should support the creation of community information bulletins that can be addressed to all individuals with relevant interests. Though much organisational communication takes place in the electronic domain, it has been argued that physical, face-to-face meetings are a more effective way to communicate ideas between individuals that have different professional backgrounds or interests. Again, in the physical world, open meetings and workshops are advertised by printed flyers; the system should support an analogous mechanism to promote these events electronically. Echoing some of the issues that Bellotti identified in her email use evaluation, NuggetMine users articulated the need for a temporal component to posted messages, requesting that it be possible to set an "*expiration date*", so that messages "*relevant only until a certain date (e.g. a conference, a local event) would disappear after that date*" [Goecks02]. However, while out-of-date messages should be cleared to minimise clutter, it was also noted that current messages should be readily available to 'late joiners' [Geyer02] – in other words, the recipients of a message should not only be determined when the message is posted, but re-evaluated continually while it remains valid.

Messages may also be used as pointers to other information sources. Fischer and Ostwald observed that pre-existing documents may be brought to community discussions to ground shared understanding [Fischer03]. Groupcast and Nuggetmine demonstrated that users will also exploit community-based KM tools as a public forum for exchange of personal "*what's cool*" messages, such as URL recommendations, that may be outwith the scope of any community involvement. One kind of individual communication can, however, be disregarded: as Koch observed, additional mechanisms for private person-to-person electronic communication are not desirable as this is adequately supported by existing telephone, email and IM options [Koch03].

4.2 Platform Requirements Framework

Awareness, community and communication are intrinsic to a cyclical process of organisational evolution. Awareness of individuals' interests gives rise to communities; in communities, interests evolve; communities communicate their members' interests – and thus the cycle repeats. Underpinning this cycle are two clear themes. Firstly, there is the principal role of the individual interest. Second, there is a powerful sense of the innovative organisation being in constant flux and a consequent requirement for timeliness, for information that is current at the present moment, which can be delivered promptly to numerous client applications. It follows that an up-to-date model of users' interests will form a strong basis of a platform to support innovation. The potential of a dynamic model of interest relationships to support a *single* innovation support function has, in fact, already been identified by Lindgren and Stenmark – they advocate an "*interest-activated*" architecture for a competence finder tool [Lindgren022] [Lindgren03]. However, if a conceptualisation of individual interests can be situated within a broader organisational representation, it will serve as the common foundation on which a comprehensive suite of innovation support tools can be constructed. With that aim, in the following sections, a framework is developed that identifies requirements of the platform's knowledge representation model, architecture and client interface.

4.2.1 Knowledge Layer

Clearly, the knowledge representation will include people and organisations. Interests, too, have been identified as a fundamental concept; in turn, it will be necessary to represent things to be interested *in*. A model of community involvement and stages of community evolution is needed. Message bulletins and social events should also be represented and, as both forms of communication typically have a specific target audience, this, too, should feature in the representation. The demand for timeliness implies that some means to define temporal constraints – and, in particular, temporal constraints relative to other temporally constrained concepts – will also be required. Relationships that must be in place between these concepts are also apparent: it must be possible to traverse a social network, or identify a Community of Practice for a given topic or interest – and, if we are to reveal competence, determine when an individual has contributed that interest to a formal group.

The necessary attributes of the knowledge layer can therefore be summarised as follows:

- Representation of fundamental organisational units – organisations themselves, groups and individual people.
- Representation of topic hierarchies that describe the domains in which the individuals modelled are professionally active.
- Representation of individuals' interests formed of combinations of these topics and of their application, through involvements, to groups.
- Representation of communications and communications opportunities – namely informal bulletins, formally published media and physical and virtual social gatherings.
- Representation of a targeting or addressing specification that can qualify individuals to which these communications and communications opportunities should be accessible, by their interests, involvements or affiliations.
- The ability to apply time constraints to those concepts where it is appropriate to do so, specified absolutely or in relation to some other temporally constrained concept.
- A class structure that is both as compact as possible and is an intuitive representation of the organisation such that it can be presented in a clear and straightforward fashion to the user.

4.2.2 Architecture Layer

At the beginning of this chapter, awareness, community and communication were identified as the three key ambitions that should be supported. From these, a number of functional components were suggested: a bulletin board, an expert finder and social network exploration tool, a community manager and an event scheduler or diary. Implicit in these requirements and the knowledge layer specification is provision to manage the underlying knowledge model; to expose and manipulate the topic trees, declared interests and personal and organisational details. The platform's architectural requirements, though, are determined not only by the functionality that it must provide, but also by the classes of device that it will service. Today, the traditional desktop PC is only one of the many different ways that individuals access computing resources. Modern work patterns see people spending increasing amounts of their working day away from their desks. They become mobile to meet [Kristoffersen99]. When the era of modern computing began in the 1960s, computing power was a scarce resource; a single mainframe would be required to serve the computing needs of many individuals and as a result access was heavily restricted. The democratisation of computing came in the late 1970s with the arrival of the personal computer, or PC and by 1984, the number of people using personal computers had surpassed those using shared computing resources [IDC96].

Now, portable, laptop and handheld PCs and, most recently, 'smart' mobile telephones have supplanted desktop machines for many tasks. Each, though, aims simply to make the desktop computing experience as portable and mobile – as *ubiquitous* – as possible. Witness the cut-down editions of standard desktop productivity software that are bundled with these platforms. *Situated* computing, in contrast, "*is based on the belief that such universality is neither attainable nor desirable*" [McCullough04]. It represents a rediscovery of the importance of place to computer interaction. Increasingly, operating alongside catch-all mobile appliances are to be found some custom pervasive computing technologies: wearable smart badges and IPADs and interactive display boards that can provide contextually sensitive ‘windows’ into a system. Primitive examples of these technologies were among the community-based KM tools surveyed.

A single individual may now have many potential points of access to the system available at any moment. The limited local computing capability of some of the devices that the system will be expected to support suggests that a thin-client / server architecture is most appropriate. The sheer potential number of clients that will be connected to the server indicates that the per-query organisational modelling and reckoning that it must perform should be limited; instead a single active process should maintain this evaluation so that it can be referenced to service any client query. An active platform will also be capable of *pushing* information, beeper-fashion, to client devices as appropriate so that the amount of speculative polling for new information that the system must service is minimised.

To summarise the key architectural and functional requirements:

- An active instantiation of the knowledge layer that provides for:
    - Maintenance of an evaluation of relationships between topics represented in the knowledge layer.
    - Maintenance of an evaluation of relationships between interests, involvements and groups.
    - Evaluation of the targeting or addressing specification.
- Driven by these evaluations, a set of components that provide:
    - A bulletin board that allows user- and machine-authored bulletins to be delivered dynamically based on the evaluation of their attached targeting or addressing specification.

- A community of interest browser or ‘expert finder’, which exploits the evaluations maintained within the knowledge representation to identify users with an interest related to another user’s declared interest or a topic in the knowledge representation.
- A group creation and management facility, that allows individual users to form ad hoc interest or activity groups whose membership requirements are declared using the targeting or addressing specification.
- A means to identify and promote awareness of groups whose overall membership involvements are relevant to the overall membership involvements of another group.
- An events diary, to facilitate organisation of face-to-face and virtual group meetings and promotion of these via the targeting or addressing specification to interested eligible attendees within the platform’s user base.
- A social network exploration tool to reveal existing social relationships and potential opportunities for social contact evident in the knowledge representation.

- Utility functions to allow exploration and manipulation of personal and organisational details in the knowledge representation.

4.2.3 Interface and Client Application Layer

The range of client devices and operating systems that must be supported also has implications for the interface and client application component of the framework. It can be expected that some of the client devices that must be supported will have limited or intermittent connectivity, either as a direct result of their mobility, or through a need to conserve battery power. It is also likely that some devices will gain their connectivity via commercial carrier that levy per-byte data transfer charges. It is therefore important that the communication mechanism provided to interact with the system should be as lightweight as possible, both in the bandwidth that it consumes and in the processing overhead that it places on the client. This implies a need not only for a compact and efficient data transfer protocol, but also that the amount of data transferred that is spurious to the function of a client device is minimised. On the other hand, the ability of the platform to support rich multi-function clients, such as kiosks or interactive signage panels, should not be compromised. Where a client device has existing functionality, provision should be made for this to be integrated synergistically with that offered by the platform.

In addition to the associated client devices that it must support, it is also desirable that the platform be capable of interfacing with external data sources. Koch [Koch03] highlights the inconvenience experienced by users required to individually maintain multiple copies of their personal profile across the range of applications and appliances found in the workplace, finding themselves entering their personal information "*again and again*". While a unified platform of innovation support tools can minimise additional burden, it will not replace the email or Instant Messaging client, let alone the organisational HR database, personal mobile phone configuration or multitude of other tools, applications and devices that require access to personal information. Aside from the inconvenience caused, it is likely that, over time, inconsistencies between the multiple profiles will arise.

Though the semantic web community has devoted some effort to the development of intelligent technologies that will identify and reconcile some such discrepancies, these operate retrospectively on large data corpuses, inevitably provide a 'best guess' solution and are generally considered an interim solution for applications that must use multiple contradictory legacy information sources. Maintenance issues notwithstanding, it is questionable whether an individual should even be 'allowed' to assert the many items of personal information of which they do not have direct ownership or control. Should, for example, they be free to present a colleague's email address, phone number or office location as their own, or, indeed, specify the post code for their building when these codes are managed by the postal authority? Unaddressed, these issues would impinge on the practical viability of our platform. It is clearly important that whenever possible, local references to non-application-specific information, *such as* an individual's email address, phone number or office location, can be coupled to definitive external sources.

The interface and client layer, therefore, requires:

- Use of a lightweight, client-platform-agnostic communication protocol that can be implemented using resource-constrained devices.
- Minimal transfer of spurious information; the platform should return only information that is needed to satisfy the client's request, with any pre-filtering necessary performed by the platform rather than the client.
- Provision of access to individual functional components, to support simple single-function clients.
- Support for daisy-chaining of client requests to enable complex client applications to be constructed that exploit multiple aspects of the platform's functionality – i.e. where possible, information transmitted to the client should include identifiers that can be used as queries or parameters for other client requests.
- The ability to 'push' information updates to clients in response to changes in the knowledge model or knowledge model evaluation, to avoid the need for them to repeatedly poll the server.

4.2.4 Framework Summary

This framework defines a set of requirements for the underlying knowledge representation, aspects of the architecture and functionality and the way that the platform should interface with the client devices that it will support and with external information sources. The remainder of this chapter develops these requirements to specify the knowledge representation model and platform architecture that will be used.

4.3 On Interest and Expertise

Inadequate representation of user interest was, ironically, a failing common to many of the innovation-support applications surveyed in the previous chapter. If the interest-activated approach we have determined to use is to be successful, an effective mechanism must be found by which the system can identify, evaluate and compare users' interests. One way to achieve this is to simply allow users to specify their interests by hand, selecting combinations of topics from a common taxonomy maintained in the system. This is undoubtedly the most straightforward approach, though it has been observed that individuals tend only to explicitly declare their peripheral interests, omitting those that relate to their day-to-day work, with the result that true experts in a subject are never identifiable to the system [Lindgren03].

Several techniques have been developed that purport to model an individual's interests automatically from an aspect of their existing computer use. Some of these techniques have too high a degree of inherent latency to be of use in an innovation-led environment; we can discount, for example, those that rely on monitoring readership of formally published documents or bookstore purchasing trends. More appropriate are techniques that infer interest from personal notes or informal inter-personal communication. Email is recognised as the primary means of inter-personal information exchange within communities [Ducheneaut00]. As exemplified by the hebb application [Carter03] outlined in Chapter 3, analysis of email has become the focus of attempts to achieve responsive automated interest recognition.

Schwartz and Wood pioneered the identification of interest groups from email archives [Schwartz92]. They developed an algorithm to identify clusters of tightly connected individuals from their email communication patterns; examining weeks of mail server logs to construct a social graph of who emailed who. The technique allows users to "*search for people by requesting a list of people whose interests are similar to several people known to have the interest in question*". It requires there to be known, identifiable, "*distinguished person*(s)" for a particular interest that will form the root point(s) of the social graph analysis. The network revealed is somewhat akin to a Community of Practice; it is a clearly identifiable group of individuals who communicate regularly over a period of time. The resulting overview of the community could raise awareness between members on opposite sides of its social graph. Messages could be addressed – or events advertised – to '*people who're into the kind of stuff that John and Tom are doing*'. This technique cannot put a name to that 'stuff' though – the only properties of an email message considered in the analysis are the addresses listed in its 'from' and 'to' fields. Moreover, it cannot recognise lone individuals or small groups that have had no contact with a named root person, or locate those sharing a novel interest for which there are yet no 'distinguished persons'.

More recently, Natural Language Processing (NLP) has been applied to mine the *content* of email messages for significant keyword clusters. Unfortunately, email presents a challenging environment for NLP because messages often are very short and assume a high degree of pre-existing common ground; "*email conversations are grounded in sufficient mutual understanding to allow very brief, sketchy and implicit references to succeed without posing significant problems in interpretation*" [Ducheneaut02]. Without access to this grounding, significant but niche or obliquely referenced concepts in the message can be missed, or an unacceptably high number of spurious concepts identified (*users of the hebb system reported that the word 'notes' was interpreted as a common interest*). If both the signal-to-noise ratio of interests identified and sensitivity of detection of genuine interests are to be improved, better access will be needed to the background and historical contexts (*the assumed common ground*) in which a communication takes place.

Nonetheless, it has been shown that by using NLP algorithms specifically tailored for this purpose, email message analysis can now successfully identify long standing, widely communicated interests – the kind that Lindgren suggests are less likely to be explicitly declared to the system. McArthur and Bruza [McArthur03] employed a modification of the HAL (Hyperspace Analogue to Language) [Burgess98] semantic representational model as the basis of a system to identify the "key players" involved with a project or product in an organisation. Similar functionality has been deployed commercially within Tacit's ActiveNet platform[1]:

> "*Here's how Tacit's ActiveNet product works: Employees are told their e-mail or other files will be scanned to compile profiles of their interests. They may opt out if they don't want to participate. E-mail is then scoured, with frequently used phrases plucked out and categorized by topic. Over time, topics that fall off your radar are discarded and new ones added.*"[2]

Email analysis may also offer a way to reveal related pair-wise collaborations that occur between individuals working in ostensibly unrelated fields; one can imagine, for example, that a seemingly 'crazy' idea for colour-shifting clothing that emerges in a dialogue between a fashion writer and designer might never be purposely disclosed to a formal organisational awareness platform, yet if the pair could be made aware of similar thoughts having emerged, serendipitously, within

[1] www.tacit.com

[2] www.usatoday.com/money/companies/management/2004-06-13-tacit_x.htm

an academic project developing smart textiles, they may be inspired to collaborate to develop the idea further.

In our context, however, the derivation of users' interests from their email chatter has a fundamental limitation: you can only send an email about an interest – and make that interest visible to the system – if you are already aware of someone who would share your interest, that you can send the email *to*. Individuals within an organisation who, unknown to each other, share a truly novel interest may work independently for months before deciding to speculatively publicise their efforts through email and finally becoming visible to the awareness system (*or, worse, abandon the work without ever becoming aware of there being a potential collaborator*). Clearly, a system to support innovation would, instead, aim to link such individuals at the earliest possible opportunity. In this regard, a useful development is the move towards integration of email (user to other, or user to others communication) and electronic note taking (i.e. user to self communication). Already, Pocket PC - based PDAs offer the facility to translate hand-written scribbles to machine-readable text and wirelessly transfer the resulting 'note' to a desktop Outlook email client. Bellotti et al. point to deeper future integration of email and Personal Information Management (PIM) functionality with their ThinkDoc [Bellotti02] prototype, in which emails, notes, scanned documents and other such material can be woven together to form a rich tapestry-like work package that individuals can pass between themselves via the ThinkDoc server.

Compared with sketching on the back of an envelope or jotting notes in a pad, any form of computer note-taking seems clumsy; what the PDA gains in spontaneity and portability, it loses by its cramped display and fiddly user interface. However, new Anoto[3] digital pen products are blurring the distinction between paper and screen, combining the intuitiveness of the spiral-bound notebook with the machine-processibility and immediate availability of computer-

[3] www.anotofunctionality.com

authored notes. Pens augmented with miniature cameras are used with special stationery, each page of which is pre-printed with a faint pattern of dots. These dots allow the pen to determine not only the unique identity of the sheet of paper on which it is resting, but also its exact position. Each ink mark made on the page is simultaneously recorded, electronically, within the pen. From there, the pen strokes are transferred wirelessly to the user's mobile phone, PDA or PC, where character recognition software turns them into ASCII text. Specially mapped dot patterns can be used to signify interactive interface elements called 'pidgets'– printed analogues of the digital text fields, check boxes and click buttons that are so familiar on the desktop. These could be used to, for example, mark an idea on the page 'private' or 'public', or indicate that collaboration is invited from colleagues. Digital pen technology is in its infancy; the first generation of consumer products is expensive and a degree of user hand-holding is required to assist the software interpret the captured page. Nonetheless, it is very clear that, as digital pen and NLP technologies evolve, it will become possible to directly infer users' interests from their paper notes.

Given the limitations of current interest capture methodologies and the promise that emerging technologies hold, it would be unwise to lock our system into a particular automated technique. A decision has been made, therefore, to use interest templates that are intuitive enough that they can be human-managed in the short term, but have a sufficiently simple grammar that automated interest identification technologies can be retro-coupled to the system through the existing interest profile management interface, either to replace user management of interest profiles outright, or offer outline profiles that the user can refine or reject, without significant system modifications being required. In the mean time, we can aim to minimise the risk that core interests will not be represented by ensuring that interest declarations are treated as an intrinsic component of any statement of group membership.

So what of *expertise*? It is tempting to assume that expertise is a characteristic distinct from interest, somehow evaluable and modellable in its own right. It is true, of course, that one might be expert in a subject without necessarily finding it very interesting. Lindgren et al., though, drew some interesting conclusions on expertise and its relationship with interest in innovation-led organisations from their study of worker experiences within Volvo's IT research team [Lindgren03]. They identified that, in this environment, competence or expertise are *not* formally distinct from interest, but emerge organically from the application of interests in a professional context: "*pursuing a professional interest in a corporate setting… leads to competence within that area*". Moreover, they noted that while employees valued the ability to quickly identify an expert in a particular field, it was also considered that in an "*ever changing environment… interest and commitment* [can be] *more important than formal competence*"; they explain that "*it is the intrinsic motivation that comes from personal interests that sets the limits for the organization's future*". The implication of this, then, is that the need is not for some separate representation of expertise, but for a means to identify *formally applied interests*.

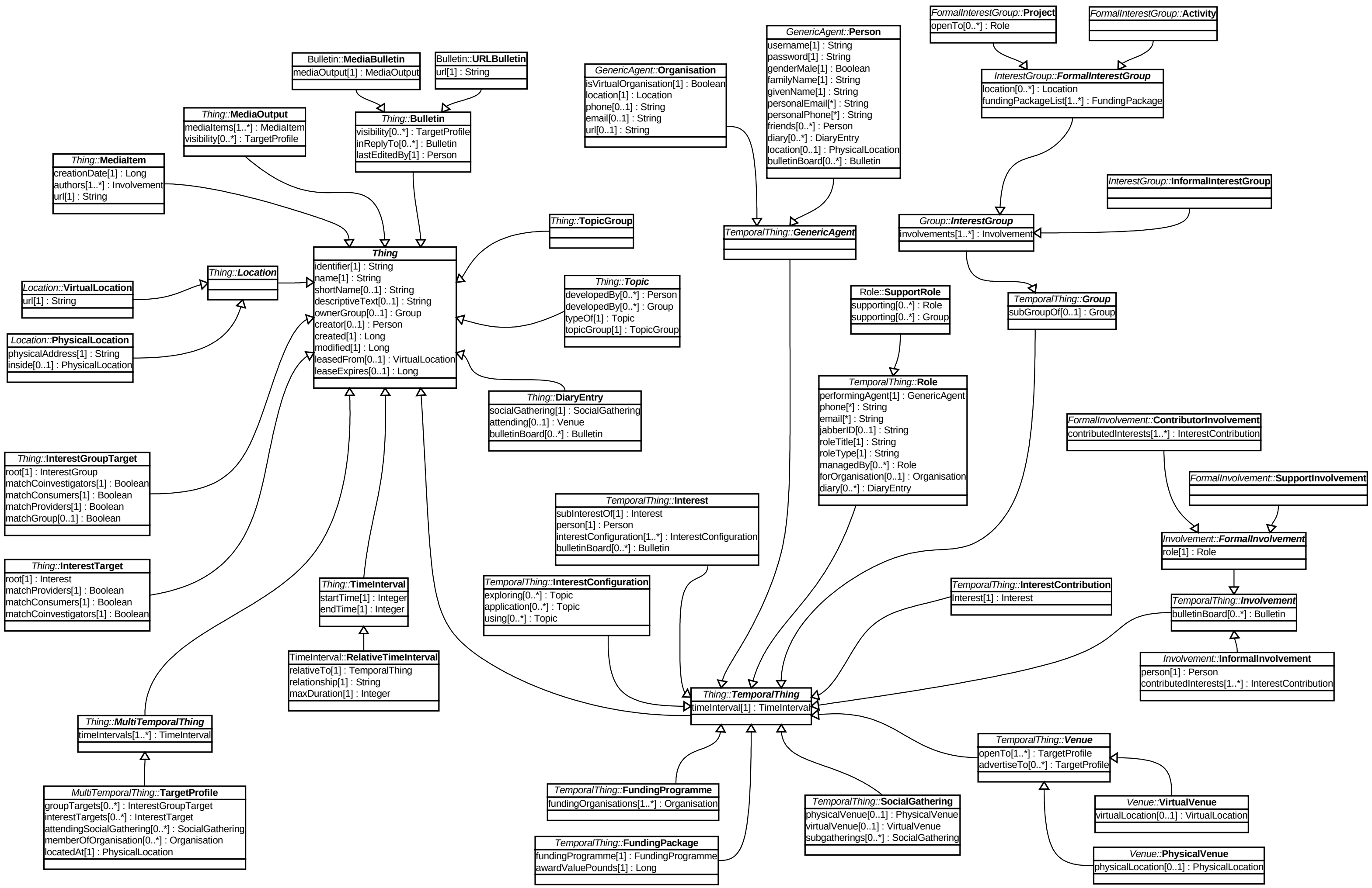
4.4 An Organisational Ontology

The framework specifies an underlying information structure through which individual interests can be situated within a broader organisational context. In Chapter 2, ontologies were introduced both as the means by which organisational structures are commonly represented within traditional IT-based KM systems and as a lightweight lingua-franca for annotation of personal information within online web pages. In contrast to the ontologies described in Chapter 2, though, our need is for an information structure that can support a single, clearly defined system – and, importantly, an *active* system. Here, an ontology must provide an *actionable* conceptualisation. Presented only with a populated copy of the ontology, the system must have enough information to decide who to show a message to, whether they might want to attend a particular event, who a user should talk to about a subject of interest and how they could best contact them. Therefore there must be no ‘gaps’ in the ontology, no missing links or assumptions of users’ background knowledge and manual intervention without which these judgements cannot be made. And, to a degree at least, its conceptualisation must be ‘warts ‘n’ all’; it should present a model of how an organisation really works, not how it wants to look like it works to the outside world. If a project group is pulling in two different directions, that fact should not be covered up in the ontology.

Additionally, the ontology must be as concise and straightforward as possible. Implicit in a framework that emphasises timeliness and ubiquity of access is a requirement that information can be presented to and interacted with by users simply and intuitively; unnecessary tiers of classification or relationship steps that must be visualised and navigated will frustrate the user and constrain the range of devices and interfaces through which the system can be accessed. Moreover, as the framework specifies an active server capable of supporting lightweight client devices rather than the dumb repository / thick client architecture commonly associated with IT-based KM systems, it is also particularly important to minimise the complexity and size of the information structure over which machine evaluation must be performed.

The schema that has been developed to satisfy these requirements is illustrated in Figure 4.1. In the following discussion of this schema, public classes will be represented in `Monospaced` type and public slots in <u>`underlinedMonospaced`</u> type. Any system-internal housekeeping or ‘utility’ slots referenced will be shown in <u>*`underlinedItalicMonospaced`*</u> type. The implemented system also employs a number of internal utility classes; where these are referenced in the text, they will be shown in *`ItalicMonospaced`* type. Neither utility classes nor the majority of utility slots are shown in Figure 4.1, as they are specific to a system’s implementation of the ontology and are not to be exported where elements of the ontology are exchanged between systems.

Figure 4.1 – An ontology for an innovation-led organisation



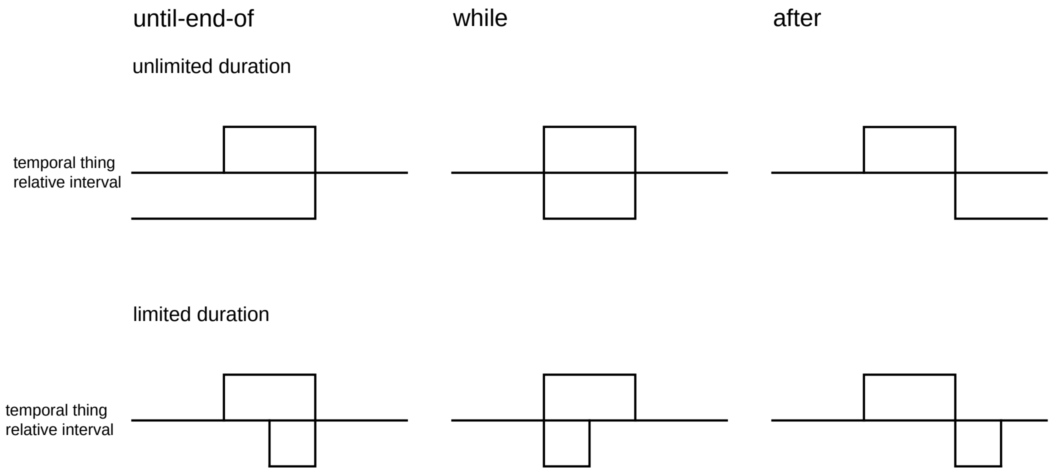


Figure 4.2 – Relative Time Interval operation

4.4.1 Temporal Dependency

Every ontology has some abstract base class from which all other classes are derived – borrowing from Cyc [Lenat90] we term this a *thing*. A `Thing` is assumed to be a constant truth; for the purposes of the system, it became valid an infinitely long time ago and will remain valid forever – or at least until the record of it is deleted. In order to represent concepts that are valid only from or to a certain point in time, a second abstract base class, `TemporalThing`, again from Cyc, is defined as a subclass of `Thing`. Clearly, a `TemporalThing` must be scoped in time. As in the AKT ontology, each `TemporalThing` references a `TimeInterval`, that specifies a <u>`startTime`</u> and <u>`endTime`</u>, during which the `TemporalThing` will be considered valid, or active, or true. A requirement has been identified though to be able to couple the validity of one `TemporalThing` to that of another. The `RelativeTimeInterval` allows simple temporal dependencies to be represented. Defined in relation to a `TemporalThing`, it supports three mutually exclusive temporal <u>`relationship`</u>s: *until-end-of*, *while* and *after*. In addition, a <u>`maximumDuration`</u> can be specified, to constrain the period of validity. The operation of this constraint is illustrated in Figure 4.2. `The TemporalThing` with which the `RelativeTimeInterval` is coupled may itself reference either a concrete `TimeInterval`, or another

`RelativeTimeInterval`. Hence, chains of temporal dependency can be established, that the system must resolve as and when the `TimeInterval` of the anchor `TemporalThing` is defined or modified. A `MultiTemporalThing` is also defined, specifically to allow messages to be scheduled for display relative to more than one event.

4.4.2 Person, Role, Organisation

The inclusion of a 'Person' class is, of course, common to all organisational ontologies. Ordinarily, though, the various job types represented are derived from this 'Person' class (e.g. [Liu02]). A typical arrangement might have, say, 'AcademicPerson' sub-classed from 'Person' and 'AcademicResearcher' subclassed from 'AcademicPerson'. Presumably such ontologies were designed for organisations with especially dedicated and focussed employees. Consider the example of the well known (and highly *innovative*) gadget design duo, Tony Dunne and Fiona Raby. Both hold academic research posts at the Royal College of Art – they would both be instances of 'AcademicResearcher'. However, they are also visiting professors at Interaction Design Institute, Ivrea and are partners in their own independent design practice, dunne+raby[4]. This combination of tenured- and freelance- or self-employment is commonplace in innovation-led organisations (*it was a feature of Bonnie Nardi's netWORK study [Nardi02]*), yet to express this situation using the ontology structure described we would need to define not two, but six instances of 'Person'. This is, of course, nonsensical; there are not three different Dunnes, rather the one 'Person', who has 'Roles' in three different 'Organisations'.

The distinction between roles and the entities that perform them was explored by Guarino [Guarino92], who suggested that unlike entities, roles are "*founded*" – a secretarial role represents secretarial support offered *by* someone *to* someone or something – and they lack "*semantic rigidity*" – when the person

[4] www.dunneandraby.co.uk

ceases to be a secretary, they remain a person. Fan et al. [Fan01] expand on this distinction, offering three characteristics that have informed the ontological structure adopted for this work:

- "*Roles are created and destroyed dynamically*" – the role may cease to exist independent of the entity that is performing it; a person does not cease to exist when the role that they are performing ends.
- "*A role can be transferred between entities*" – for example, a role within a group or organisation performed by a person might survive the departure of that person.
- "*An entity may play different roles simultaneously*" – in the example above, one person performs roles in three different organisations at once.

Hence in our ontology, roles are represented independent of the entities, the `GenericAgent`s, that perform them. One `GenericAgent` – typically, in this application domain, a `Person` – can perform any number of `Role` instances, each one of which can optionally be owned by an `Organisation`. Attributes of the `Role` such as <u>`email`</u> and <u>`phone`</u> may mirror those of the performing agent, or persist when the `Role` is transferred from one agent to another. A subclass of `Role`, `SupportRole`, is also provided. For the purposes of this system, this signifies simply that in that `Role` the individual is not to be considered intellectually engaged with the `Organisation`. Another noteworthy feature of the `Person` class is a <u>`friends`</u> slot; an idea borrowed from FOAF's 'knows' relationship, this allows an individual to specify social contacts independent of any formal professional involvement.

4.4.3 Interests and Topics

While an interest-activated approach depends on different individuals having similar, or at least related, interests, a person's interests are, ultimately, *personal* to them. They are shaped by their own, unique, professional and life experiences. The building blocks of these interests, though – the things that the interests are about – are usually much more widely held; recall Drucker's comment that innovation is the application "*of knowledge to tasks that are new and different*" [Drucker92]. (*If this were not the case, individuals would find communicating their interests to others extremely difficult*.) Accordingly, in this ontology, an `Interest` is considered to be an individually owned, personal interpretation of a collection of commonly understood taxa – given the importance placed on conversation and exchange of ideas, and echoing the Semantic Web Portal Ontology, we call them `Topic`s. This personal interpretation must be captured in a way that allows meaningful interest comparisons to be made, yet is simple enough that our requirement that interest profiles be manually manageable is satisfied. A scan of organisational websites reveals that individuals' interest descriptions consist, essentially, of three concept lists: the primary focal concepts that they are *exploring*, concepts that describe the intended *application* of their exploration and concepts that they are *using* to support their exploration. This structure is represented in our ontology. So that the gradual evolution of an individual interest over time can be modelled, these concept lists are packaged as temporally-windowed `InterestConfiguration`s, any number of which can ultimately be attached to a single `Interest` to represent its development.

The richness of interest relationship that can be inferred depends in part on the complexity of the underlying topic structure. In this application, it is desirable that the system be able to apply generalised interest relationship rules that might not be expressly represented in two particular interest definitions. Figure 4.3 depicts a section of `Topic` tree that users might construct if the system were deployed in a computer science research environment. Two types of inter-`Topic` relationship are apparent in this example. There are, of course, direct familial relationships: Bluetooth is relevant to Wireless Networking because Bluetooth is a child of, or a *type* of Wireless Networking. We observe, though, that there is an associative relationship too: Bluetooth is also relevant to Mobile Computing because Mobile Computing is commonly *supported by* Bluetooth. The `Topic` class specified in the ontology allows both familial sub-classing and these cross-concept support relationships to be declared. As the system is intended to support any kind of innovation-led organisation, no predefined `Topic` types are included in the ontology, though individual system deployments may subscribe to commonly managed `Topic` trees (taxonomies). A single `Interest` profile can reference `Topic`s from different taxonomies.

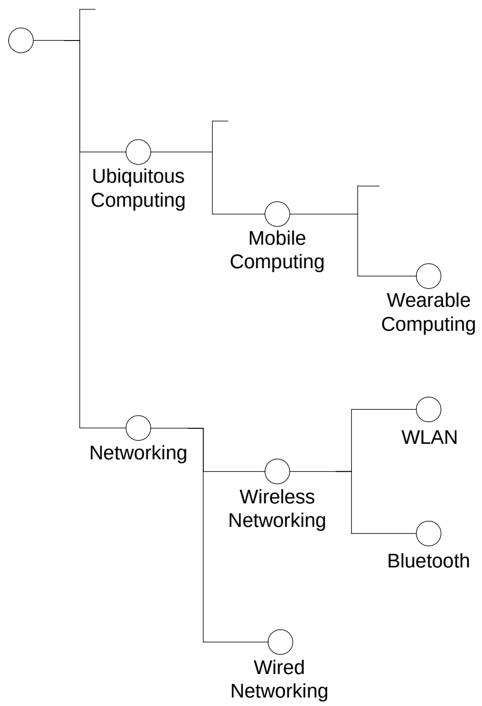


Figure 4.3 – Partial topic tree

4.4.4 Communities of Interest

The term ‘Community of Interest’ has been seen to encompass a series of evolutionary stages, from informal coalition building to formal, funded project.

The initial coalition building stage is represented by the InformalInterestGroup class, an open community with a fluid membership of individuals that wish to explore a mutual interest. The FormalInterestGroup abstract class represents a progression to a closed, task-oriented community[5]. Project membership is essentially fixed; individuals who have been personally invited to join are listed in the openTo slot so that a client application will allow them to create an Involvement with the community. A location slot is added, which can be used to state physical or virtual locations where the community can be found. A fundingPackageList slot is also included, through which sponsors can be referenced. Individual packages of work, either within or independent of a Project, are represented by the Activity class.

The involvement group classes do not accommodate explicit structured representations of their overall sphere of interest (*though a slot for a natural language text description is inherited from the Thing class*). Instead, the representation of Communities of Interest membership reflects that individual members bring different interests to the community. For FormalInterestGroup-derived classes, the Interests that a Person contributes in a single Role are packaged as a ContributorInvolvement fact, which serves as a declaration of membership[6]. This treatment of involvements as rich first-class objects again contrasts with pre-existing organisational ontologies where community involvement is recorded only as a semantically impoverished membership list slot on the community class; a simple list of person identifiers that tells nothing of the *nature* of any individual's involvement. For InformalInterestGroups, the InformalInvolvement class allows an individual to associate themselves personally with the community outwith any

[5] The FormalInterestGroup classes can also be used to represent 'groups of one' – personal formal work packages that do not involve collaboration with other individuals.

[6] A SupportInvolvement class, in which no Interests are tabled, is also provided to allow secretarial or administrative support for a CoI to be modelled.

current work role. By examining current associated `Involvements`, the system can assemble a real-time overview of the community – *and identify formally applied interests* – without community members having to negotiate and maintain a single overall model of their activities; the requirement that individuals be encouraged to declare their core interests to the system is also met.

4.4.5 Communication: Bulletins, Documents and Gatherings

A bulletin metaphor has been suggested to raise awareness of communities and events and allow exchange of website and document recommendations. A need has been identified to ensure that these bulletins are made available to all potentially interested individuals but discarded automatically when they are no longer relevant. The ontology satisfies these requirements by pairing `Bulletin` classes, that convey the message content, with the rich, dynamic addressing capability provided by `TargetProfiles`, that determines not only the individual recipients of the `Bulletin`, but also the time period during which it should be visible.

Bulletins are actively brought to our attention, *pushed* into our consciousness by being posted on bulletin boards. The products of community work, such as documents and slideshow presentations, are not usually systematically push-published in this way. To use another physical metaphor, these are more likely to be deposited in a library to be discovered by individuals over a period of time. Reflecting this distinction, media items are represented separately in the ontology. A `MediaItem` fact associates one or more community `Involvements` with the URL of the piece of work. The `MediaOutput` class forms a targeted instance of publication of one or more `MediaItems`. By separating the `MediaItem` and its publication, the `MediaOutput`, the ontology supports multiple publication rules. One can conceive, for example, that an individual might be free to publish documents that they have authored as they choose within their organisation, but that the decision of whether to present the work externally might be taken by a publications committee.

Though MediaItems are not routinely actively promoted by their authors, they *are* often brought to the attention of others by their readers in an ad hoc manner, for discussion in Communities of Interest, or simply as recommended reading to colleagues with related interests. The MediaBulletin class allows targeted recommendations of document publications to be made. Because each MediaBulletin concerning the same document will, indirectly, reference the same MediaItem fact, a client application can consolidate presentation of multiple recommendations in a single notice to the user. A URLBulletin class is also provided, allowing web addresses to be similarly recommended, though these are referenced directly as it cannot be assumed that there will be locally managed meta-descriptions for external websites.

Another form of communication that the ontology must accommodate is the social gathering or meeting. As it is now commonplace for small, physically private meetings to be made available to a wider audience via video-conferencing or streaming technology, a model has been adopted that allows multiple Venues to be attached to a single SocialGathering, with each having its own access criteria. Recognising that a SocialGathering may be open to a wider group of people than are actively encouraged to attend, the Venue class includes two target slots – <u>openTo</u>, signifying all individuals that are permitted to join the gathering and <u>advertiseTo</u>, signifying those to which the event should be promoted. (*By default, both slots would reference the same* TargetProfile*(s).*) Individual interest in, or planned attendance of, a gathering is conveyed in the DiaryEntry class, that links a Person (and, optionally a Role) instance to a SocialGathering instance.

4.4.6 Targeting People

Delivery of bulletins, access to published media and eligibility to attend a social gathering or join a Community of Interest are all governed by `TargetProfile` instances. `TargetProfile`s are a novel contribution of this work, perhaps best considered as rich, active, 'search boxes'. A target profile is actually composed from three components. As well as the `TargetProfile` fact itself, there are two sub-target types, `InterestGroupTarget` and `InterestTarget`, that it can be linked to, to refine the overall target, or 'search query'. Each sub-target takes a root fact identifier – referencing an `InterestGroup` or `Interest` – and a number of non-exclusive Boolean options that constrain the matching facts by their relationship with the root fact. Namely, these restrict the sub-target matches to instances that are deemed to be *providers* for, *co-investigators* of, or *consumers* of the stuff represented by the root `Interest` or `InterestGroup`. Additionally, the `InterestGroupTarget` `matchGroup` Boolean slot allows members of the root `InterestGroup` itself to be targeted.

```
(defrule rule-displayMessage-involvementBulletin

    (Bulletin
          (identifier ?messageIdentifier)
          (visibility $?messageTargets)
          )

    (TargetProfile
          (identifier ?targetProfileIdentifier&:(member$ ?targetProfileIdentifier $?messageTargets))
          (current TRUE)
          (matchingInvolvements $?matchingInvolvements)
          )

    ?involvementFact<-(Involvement
          (identifier ?involvementIdentifier&:(member$ ?involvementIdentifier $?matchingInvolvements))
          (bulletinBoard $?bulletinBoard&:(not (member$ ?messageIdentifier $?bulletinBoard)))
          )
=>
    (modify ?involvementFact
          (bulletinBoard (create$ ?messageIdentifier $?bulletinBoard))
          )
)
```

Figure 4.4 – An example JESS rule

The `TargetProfile` also has slots in which physical location, organisational membership and gathering attendance constraints can be specified; profile matches can be limited to individuals that are currently located at a particular `PhysicalLocation`, are members of a specified `Organisation` or are attending (*or have attended*) a specified `SocialGathering`. Exactly how target matches are identified and the targets populated is determined by the system that implements the ontology. One such implementation is detailed in the next chapter. Because the `TargetProfile` class is a subclass of `MultiTemporalThing`, the temporal constraints described in Section 4.5.1 can also be applied. This allows, for example, a `Bulletin` to be displayed to individuals with a particular interest *during* a specified event. As multiple `TargetProfile`s can be associated with a single `Bulletin`, `SocialGathering`, `InformalInterestGroup` or `MediaOutput`, they provide a very flexible targeting mechanism.

4.5 Active, by Inference

In itself, the ontology is merely a static representation of the organisation at a moment in time. To facilitate the dynamic organisational model that has been specified, the ontology must be held within an active system. The relationships expressed in ontologies facilitate machine reasoning, or inference, which forms the basis of intelligent, informed decision making. Machine reasoning is performed by *reasoning engines*. Reasoning engines parse a body of data using sets of instructions or *rules*, which stipulate the desired reasoning behaviour. Each rule comprises an *if* portion and a *then* portion. The *if* portion – the Left Hand Side – defines a set of patterns, and the *then* portion – the Right Hand Side – defines some actions which will be executed if the patterns in the *if* portion are *matched*. This process, which continues until all satisfied rules have been executed, is called *pattern matching*.

The reasoning engine selected for this work was the *Java Expert System Shell* [Friedman99] (JESS), developed by Sandia National Laboratories. A Java variant of NASA's C Language Integrated Production System (CLIPS), it was chosen because of its minimal system requirements, scalability, ease of integration with external code and outstanding quality of developer support. An example

JESS rule taken from the implemented system is shown in Figure 4.4. The rule, to post a `Bulletin`, is triggered by a relationship between three facts in the system's data model: the `Bulletin` itself, a `TargetProfile` referenced by the `Bulletin` and an `Involvement`, the *bulletinBoard* slot of which the Bulletin should be posted to if it has not already been. The *then* portion of the rule, which follows the '=>' symbol, acts on the data model to add the `Bulletin`'s `identifier` to the *bulletinBoard* slot of the `Involvement`.

A practical rule-based system might contain several hundred such rules. If the reasoning engine were simply to re-evaluate each from scratch against the entire data model in a continuous cycle, it would be extremely inefficient. JESS instead implements an optimised pattern matching algorithm called *Rete* [Forgy82]:

> "The Rete algorithm is implemented by building a network of nodes, each of which represents one ore more tests found on a rule LHS. Facts that are being added to or removed from the working memory are processed by this network of nodes. At the bottom of the network are nodes representing individual rules. When a set of facts filters all the way down to the bottom of the network, it has passed all the tests on the LHS of a particular rule"

Like ball bearings in a child's toy, accumulations of facts that have partially satisfied a rule pass down the network, test by test, packaged as tokens. Because this record of partial rule satisfaction is maintained, when a fact is added to or retracted from the data model[7], only this fact need be matched against the rules. By using this optimisation, JESS is able to support real-time interaction with very large data models.

[7] Modification of an existing fact in the data model is achieved by *retracting* the old fact, then *asserting* a new fact that reflects that desired changes.

The action triggered by a rule might be a modification of the data model, which could itself cause further rules to be satisfied. Alternatively, it might be a call to an external Java function (*and, therefore, potentially to another application or device*). A JESS-based KM system can therefore push information proactively to client applications and devices in response to changes in the populated organisational ontology. A rule could, for example, notify an individual when a bulletin becomes visible to them by calling a Java function that messages a mobile device that they are carrying. The device need not repeatedly poll the system for updates, it can just 'wait to be told'. On occasion, though, users will wish to pull information from the system at a time of their choosing; *ask* it stuff, not be told. JESS satisfies this requirement in supporting stored queries – essentially, rules that have no *then* portion. A query returns a list of all facts, or all combinations of facts, that satisfy a set of conditions; one might be used to return all the bulletins that are visible to the user at the present moment. The ability to responsively support both push and pull interaction around a single large-scale data model makes JESS an ideal choice to form the basis of an interactive community-KM system.

Having emphasised the importance of exploiting available external information sources, some consideration must be given to how their contents can be made available to the reasoning engine. It is necessary that the information model held within the reasoning engine is synchronised with external sources, but also that these sources should not represent multiple potential points of failure; the system should not be dependent on their availability to function. Borrowing from ideas well-developed in the field of computer networks and inspired by the broader requirement for Semantic Web tools better suited for deployment in distributed and pervasive computing environments, provision will be made for information to be retrieved from external sources on a *lease* basis. When a piece of information is requested, it would be returned together with a value representing its period of validity; the transaction would be logged by the server. Should the information become invalid – that is, be modified or retracted – before the lease offered expires, the server undertakes to transmit this change to the client. A suitable middleware would ensure that this change could be propagated via intermittent connections. When the lease period expires, the server erases the transaction record. At this point, the client must re-request the item of information to be sure that its copy remains current. The server would be free to vary the lease periods offered depending on the frequency with which the item of information was expected to change, thereby allowing the size of transaction table and amount of client queries to be balanced.

Each item of information in circulation will be associated with a definitive host source; this might be a multi-processor state- or corporate-run server, or a lightweight personal mobile device along the lines of Intel's Personal Server [Want02]. Obviously low-powered, intermittently available devices cannot service large volumes of queries, so it is not envisaged that client applications will directly query primary information sources. Instead, primary sources will lease their information to a number of publicly accessible secondary triple stores, which in turn will allow queries from tertiary triple stores and client applications. A similar multi-tiered architecture is successfully employed to propagate accurate time from primary atomic sources to public NTP servers. Hence the approach described would allow a far wider range of information sources to contribute directly to the Semantic Web than is currently possible.

4.6 Architectural Overview

Figure 4.5 shows the outline architecture of the innovation support platform, Aware, that has been developed. At its centre is a single JESS instance that implements the ontology described and holds all the facts and relationships that describe the organisation. It is loaded with 310 rules and stored queries that define the system behaviour; the most important of these are detailed in the next chapter. This reasoning engine package is wrapped in a Java HTTP server that presents a lightweight XML interface for client applications via TCP/IP sockets. Because the responsiveness of the system is directly dependent on the ability of

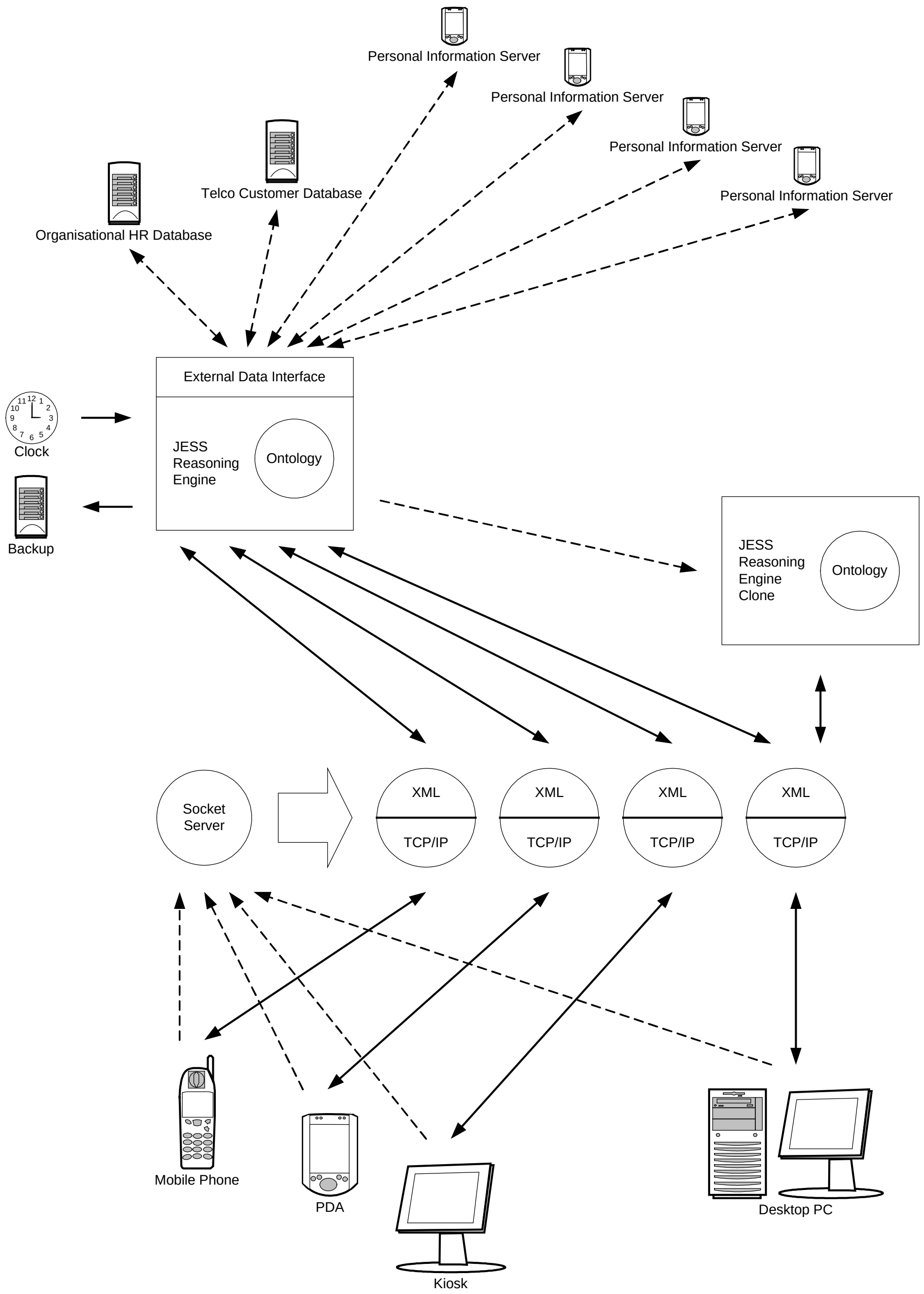


Figure 4.5 – Aware System Outline Architecture

this primary reasoning engine to service user queries as they arrive, complex atomic operations carried out by this engine at the behest of one user would severely degrade the experience of others. To address this, a mechanism is provided to generate per-connection, low-priority, read-only clones of the reasoning engine that can be used to service those user queries that may be computationally intensive. The Social Network identification function, discussed in section 5.2.2, is made possible by this mechanism. Additional utility services include a 5-minute interval clock that increments the reasoning engine's temporal representations, a background disk backup service and a stub interface for future integration with lease-principle information sources.

4.7 Summary

Awareness, community and communication have been identified as the primary requisites of innovation. Two concepts, individual interest and temporal sensitivity, have been shown to underlie each of these requisites and an interest-activated approach has been advocated as the basis of a unified framework for innovation support. To support this, an organisational conceptualisation or ontology has been developed that situates a dynamic interest model within a broader network of organisational relationships. A software architecture has been outlined in which an instantiation of the ontology is maintained within the JESS reasoning engine, forming an active core. In the following chapter, the implementation of this core will be described in detail.

Chapter 5

Implementation of the Aware Core

This chapter describes the implementation of the active core of the Aware platform. The discussion is broken into two sections: the first considers the rule-based, data-driven mechanisms that automatically maintain the basic networks of relationships and dependencies in the fact base, while the second part examines some of the on-demand query and inferencing services that client applications can invoke to mine these relationships.

5.1 Core Overview

The organisational conceptualisation, the populated ontology, forms a common bus around which the various functional components of the core are arranged. This is illustrated by Figure 5.1. Components that are rule-initiated, that is, those that react to conditions within the organisational conceptualisation, are coloured orange in the diagram, while those that respond directly to instruction from client applications are shown in blue.

5.2 Rule-Initiated Operation

The most fundamental requirement of this system is an efficient mechanism to dynamically track and evaluate pair-wise relationships between the many user-managed `Topic`s, `Interest`s and `InterestGroup`s represented in the JESS fact-base. In the following sections we will discuss the pair evaluation

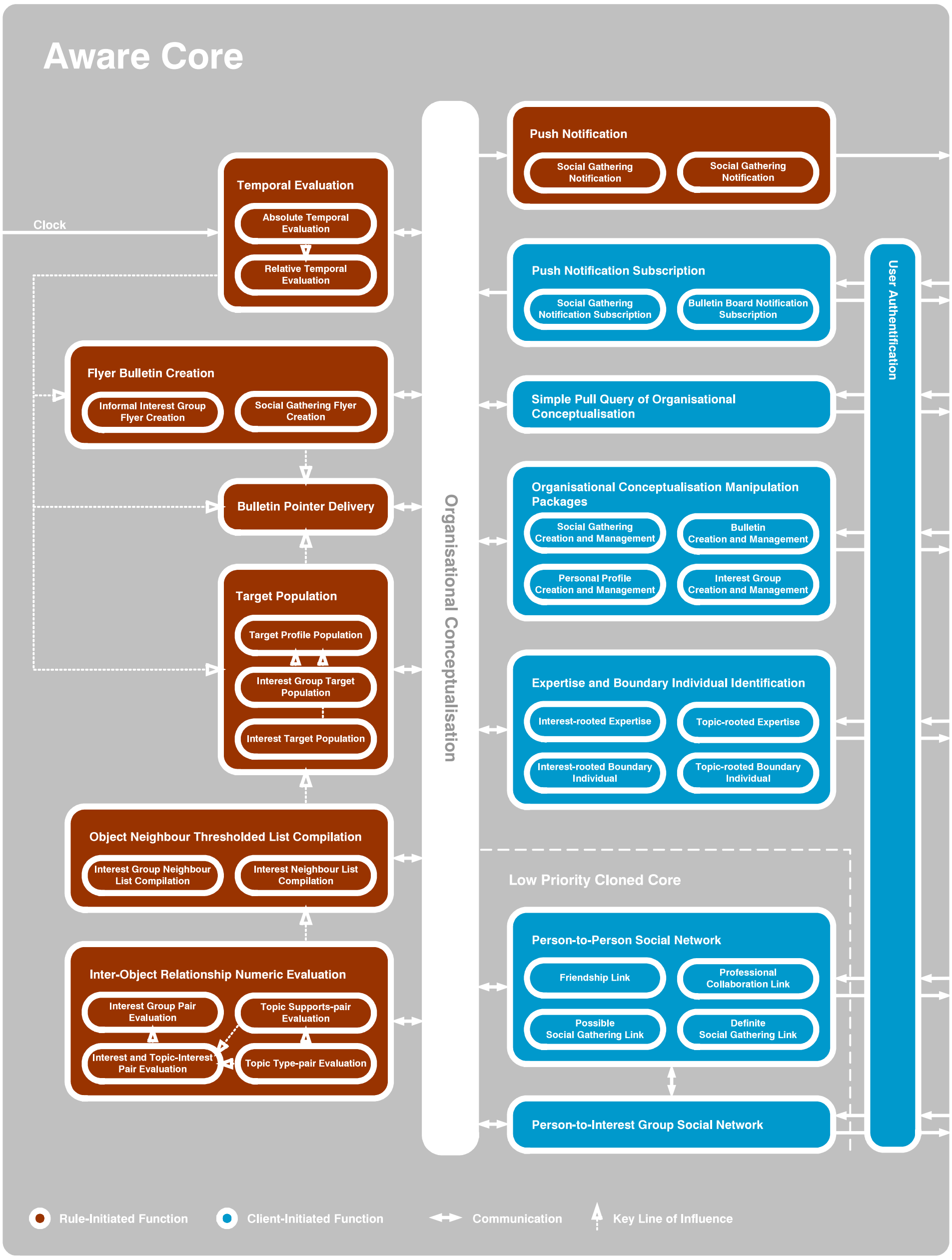


Figure 5.1 – Aware Core Architecture

algorithms employed to populate the *TopicPairScores*, *InterestPairScores* and *InterestGroupPairScores* that record these relationships in the fact-base and the evaluation trigger rules that have been developed that ensure that changes to the relationships are accurately reflected. The intention is not to make definitive, 'scientifically accurate' relationship judgements, but instead to exploit some commonsense rules-of-thumb that will provide a estimates of relevancies of `Topics`, `Interests` and

InterestGroups to each other, in the context of the organisation in which the system is deployed, that are acceptable to and predictable by the user.

5.2.1 Topic-Pair Relevance Scoring

The topic-pair relationship is arguably the most important in the system; it is the keystone upon which relevancy judgements between Interests and InterestGroups are based and, as such, is indirectly responsible for most of the rule-based actions that the system performs. The Topic fact allows two forms of immediate Topic relationship to be declared. Initially, it was intended that both support and type or class relationships would be specified manually as part of the user-authored Topic declaration. Inevitably, though, this would carry the risk of reinforcing preconceived or 'conventional' views of Topic deployment – something that is very much at odds with the stated intention to foster innovation. The need to support visualisation and maintenance of two dimensions of Topic relationships would complicate client user interface development and introduce an additional overhead for the user. It was also apparent that if future interoperation *between* systems was to be accommodated, it would be necessary for systems to subscribe to common Topic ontologies while remaining sensitive to locally emerging relationships. It is important, too, that *multiple* Topic ontologies that represent diverse domains can be supported in a single system deployment, in combinations that their authors had not envisaged. It was therefore decided that while Topic *type* relationships would be explicitly stated, *support* relationships should be derived automatically from evidence extant within the local system fact-base. The *TopicPairScores* fact, then, has two components, one derived from absolute and relative positions of the Topics within the overall user-managed Topic tree, the other from potential Topic associations inferred from the combinations of Topics represented in recent user Interest profiles.

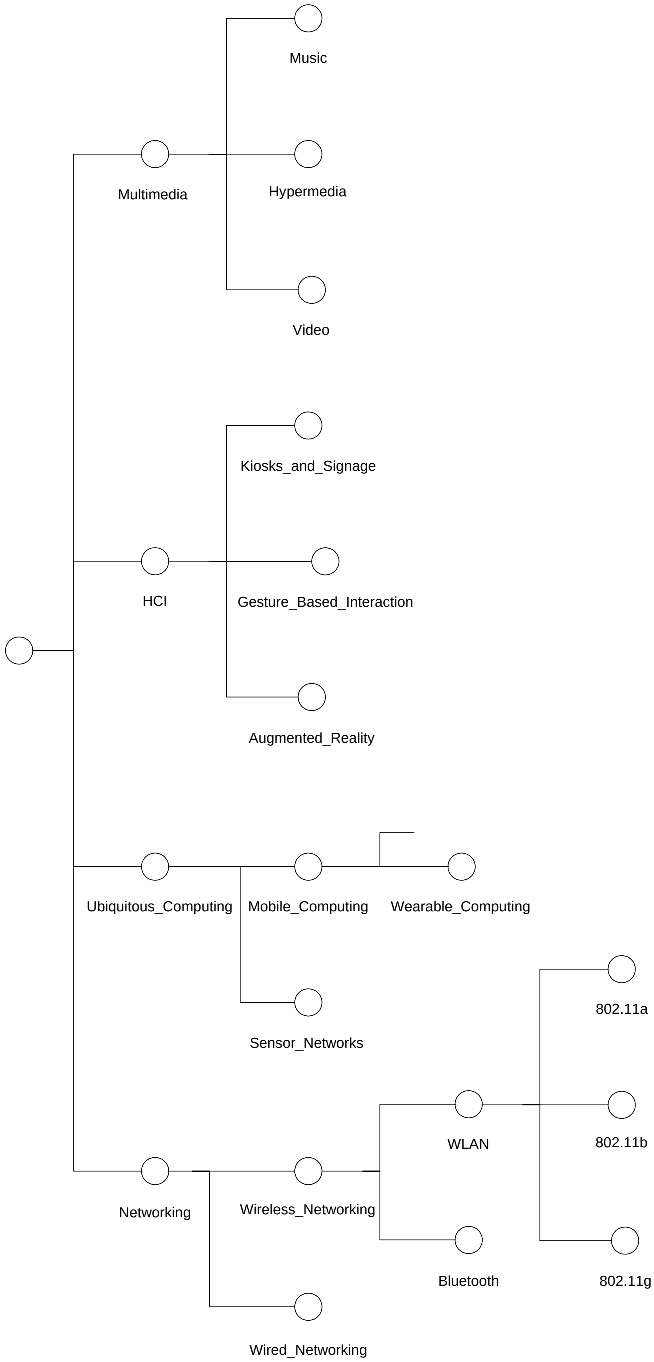


Figure 5.2 – Partial topic tree

The estimate of type-relevance of one tree-adjacent `Topic` to another is based on two properties of a topic tree. Firstly, each tier of topics represents refinements or specialisations of the previous tier; in the example Computer Science tree shown in Figure 5.2, '*Mobile_Computing*' represents a refinement of '*Ubiquitous_Computing*' and '*Wearable_Computing*', in turn, a refinement of

‘*Mobile_Computing*’. Therefore, the relevance to a topic of progressive tiers of ascendants diminishes more rapidly than that of progressive tiers of descendents, because the descendents are entirely encompassed by the topic, whereas the ascendants are only partly concerned with it. Secondly, the conceptual distinction between topic tiers is greatest closest to the root; considering again the tree in Figure 5.2, ‘*802.11b*’ represents a more subtle refinement of ‘*WLAN*’ than ‘*Wireless_Networking*’ does of ‘*Networking*’. Similarly, a tree of living things might have in its first tier ‘*Plant*’, ‘*Animal*’ and ‘*Virus*’, while an end tier might contain various breeds of dog; clearly the leap from ‘*Dog*’ to ‘*Alsatian*’ is substantially less than that from ‘*Living Thing*’ to ‘*Animal*’.

The relevance of a `Topic` to an ascendant `Topic` – i.e. a child-to-parent or grandchild-to-grandparent – is determined solely by an estimation of the conceptual distinction of the two `Topic`s from their absolute tier positions within the tree structure:

$$relevance_{child\_to\_parent} = \left(\frac{tier_{parent}}{tier_{child}} \right)^{V_{position}}$$

The relevance of a `Topic` to a *descendent* `Topic` – a parent-to-child or grandparent-to-grandchild relationship – is determined by the same formula with an additional scaling component to represent the more rapid rate that progressive ascending tiers loose relevance:

$$relevance_{parent\_to\_child} = \left(\frac{tier_{parent}}{tier_{child}} \right)^{V_{position}} \div \left(tier_{child} - \left(tier_{parent} - 1 \right) \right)^{V_{span}}$$

Variables $V_{position}$ and V_{span} determine how pronounced the effects of the two characteristics will be on the relevance estimation; the default values used by the system are $V_{position}$=0.5, V_{span}=0.5 – increasing either factor increases the influence of the corresponding characteristic on the relevance evaluation. Figure 5.3 illustrates type relevance evaluations of selected topic pairs from Figure 5.2, performed using the default values. The effect of changing these variables can be seen if the relationships between ‘*Wireless_Networking*’ and ‘*802.11a*’ are re-

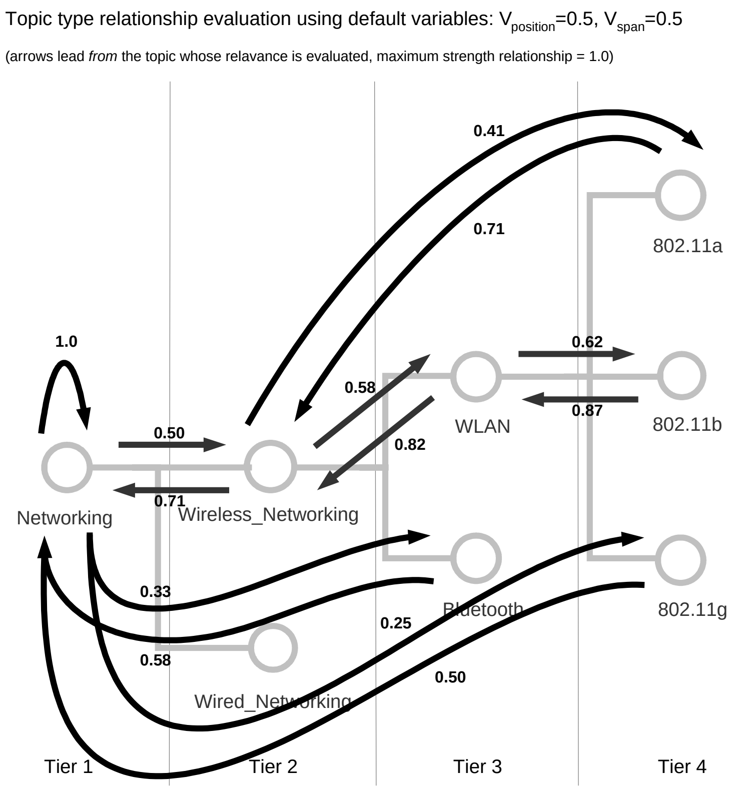


Figure 5.3 – Topic type relevance evaluation

evaluated, firstly with $V_{position}$ increased to 1, and secondly with V_{span} increased to 0.7:

$$relevance_{child\_to\_parent} = \left(\frac{tier_{parent}}{tier_{child}}\right)^{V_{position}}$$

$$= \left(\frac{2}{4}\right)^{1}$$

$$= 0.50 \quad \text{(original value} = 0.71)$$

$$relevance_{parent\_to\_child} = \left(\frac{tier_{parent}}{tier_{child}}\right)^{V_{position}} \div \left(tier_{child} - \left(tier_{parent} - 1\right)\right)^{V_{span}}$$

$$= \left(\frac{2}{4}\right)^{0.5} \div \left(4 - (2 - 1)\right)^{0.7}$$

$$= 0.33 \quad \text{(original value} = 0.41)$$

The associative *supports* and *supportedBy* Topic-pair relevancies also maintained in the *TopicPairScores* fact are obtained by examining the Topics referenced in users' Interest profiles. To ensure that recent Topic usage trends are not masked by obscure Topic applications by individual users or historic applications that might no longer be relevant, consideration is only given to those Interests currently contributed to FormalInterestGroups. It is considered, then, that '$Topic_1$' supports '$Topic_2$' if a declared Interest, currently contributed to a FormalInterestGroup, has an attached current InterestConfiguration stating that it is either:

using '$Topic_1$' or a descendent topic of '$Topic_1$'

and exploring '$Topic_2$' or a descendent topic of '$Topic_2$'

or

using '*Topic1*' or a descendent topic of '*Topic2*'

and has application '$Topic_2$' or a descendent topic of '$Topic_2$'

Again referring to the topic tree in Figure 5.2, consider this example: Tom declares an Interest entitled '*Sensor Nets*', with the InterestConfiguration ((using '*Bluetooth*') (exploring '*Sensor_Networks*')). He contributes this Interest to a FormalInterestGroup, '*Smart Environments*'. From this, it can be inferred that:

- '*Bluetooth*' supports '*Sensor_Networks*'
- '*Bluetooth*' supports '*Ubiquitous_Computing*'
- '*Wireless_Networking*' supports '*Sensor_Networks*'
- '*Wireless_Networking*' supports '*Ubiquitous_Computing*'
- '*Networking*' supports '*Sensor_Networks*'
- '*Networking*' supports '*Ubiquitous_Computing*'

However, though factually correct, it is not necessarily the case that every one of these support relationships is significant within the context of the organisation in which the platform is deployed. If the majority of declared interests `using` '*Networking*' had no relation to '*Sensor_Networks*', for example – in other words, if '*Sensor_Networks*' was only one of a great many different applications of some kind of '*Networking*' within the organisation – it could not be assumed that someone with a declared interest in '*Networking*' would wish to be deluged with messages about '*Sensor_Networks*'. Therefore to achieve an estimation of the strength of support that one `Topic`, '$Topic_1$' offers another, '$Topic_2$', the following steps are taken:

1. All `InterestConfiguration`s matching the criteria stated above are retrieved.
2. A list of all `Topic`s that these `InterestConfiguration`s are `using` is assembled.
3. If '$Topic_1$' is contained within this list, a scaling factor for the support evaluation of 1.0 is recorded; otherwise the most strongly type-related descendent of '$Topic_1$' is identified and the strength of its relationship to '$Topic_1$' recorded as the scaling factor.
4. All `InterestConfiguration`s not `using` this most significant `Topic` are disregarded.
5. For each remaining `InterestConfiguration`, the average descendent type-relevances of `Topic`s that are listed in its `exploring` and `application` slots to 'Topic2' are obtained (an occurrence of '$Topic_2$' itself contributes a relevance of 1.0), and the greater of these values retained as a measure of the evidence offered by that `InterestConfiguration` for a support relationship between '$Topic_1$' and '$Topic_2$'.
6. An average of the `InterestConfiguration` evaluation values is calculated and multiplied by the scaling factor obtained in step 3, to achieve the final evaluation of the support relationship strength of '$Topic_1$' to '$Topic_2$'. The inverse *`supportedBy`* relationship, also

maintained in the *TopicPairScores* fact, is simply the reverse support relationship: $\text{relevance}_{\text{supportedBy}}$('$Topic_1$' to '$Topic_2$') = $\text{relevance}_{\text{supports}}$('$Topic_2$' to '$Topic_1$').

In the example presented here, then, if another current, formally-contributed Interest has an attached InterestConfiguration stating that it is using '*Networking*', the indirect link that Tom's Interest suggests between '*Networking*' and '*Sensor_Networks*' will be disregarded. In reality a user would not be expected to – continuing our example – immediately consider '*Networking*' irrelevant to '*Sensor_Networks*' researchers just because its first formal use within their organisation involved a video distribution application. A future refinement, therefore, would gradually diminish the influence of inherited support relationships with time and as more direct Topic applications were declared. This can be achieved without modifying the current data or rule structures.

5.2.2 Interest Pair Relevance Scoring

Having established an up-to-date network of relationships between individual Topics, it is then possible to identify the relevant pairings of users' current Interest profiles. In comparing Interests the aim is not only to identify those that are equivalent – in other words, broadly the same interest expressed by different individuals – but also reveal indirect relationships between Interests that provide or receive support from one another. Correspondingly, the *InterestPairScores* fact has three major components: *coinvestigatorStrength* that reflects direct similarity, *providerStrength* that indicates the degree to which $Interest_{OF}$ would be considered to support $Interest_{TO}$ and *consumerStrength*, a measure of the support $Interest_{OF}$ would be considered to receive from $Interest_{TO}$. It is necessary to maintain both provider and consumer values for each pair because inter-Interest support relationship scores are typically asymmetrical. This statement might seem nonsensical – how can *X* strongly support *Y* and yet *Y* not be strongly supported by *X*? It is important to recognise that the numerical scores obtained do not represent value judgements of the contribution that one Interest makes to another. Rather, they are estimates of how relevant, or interesting, an individual who has declared one `Interest` (*Interest*$_{TO}$) would regard a message, event or group related to another `Interest` (*Interest*$_{OF}$) declared by a colleague; it can be expected that an individual whose `Interest` has a single, tightly focussed application will be keen to follow developments in that area, but if the application receives support from a wide range of `Interest`s, we would not automatically presume that an individual working in this area would wish to closely follow each of those `Interest`s, unless they had explicitly stated this by declaring additional `Interest` profiles.

`Interest` pairs are determined by the relationship between their current `InterestConfiguration`s. The `InterestConfiguration` slot combinations considered to be evidential of consumer, co investigator and provider pairings are illustrated in Figures 5.4 – 5.6. These form the basis of the rules that trigger *`InterestPairScores`* evaluation. `Interest` pairs will typically be linked by only one of the three relationship classes, so provider, co investigator and consumer pair evaluations are triggered and processed individually. Each of the relationship classes is indicated by a number of different combination templates; in Figure 5.5, for example, it can be seen that there are three basic combinations of `InterestConfiguration` that indicate a coinvestigator relationship. The strength of an `Interest` pair relationship is determined by the strength of the familial Topic relationship that satisfies the corresponding template. However, as shown in Figures 5.4 – 5.6, there are two or more `InterestConfiguration` combination templates that indicate a particular relationship. In addition, a pair of `InterestConfiguration`s might satisfy one of these templates with more than one combination of `Topic` references. Therefore where more than one `Topic` pair relationship indicates a particular class of `Interest` pair, the most highly scored `Topic` pair determines the strength of the `Interest` pair relationship. An example consumer pair evaluation is shown in Figure 5.7.

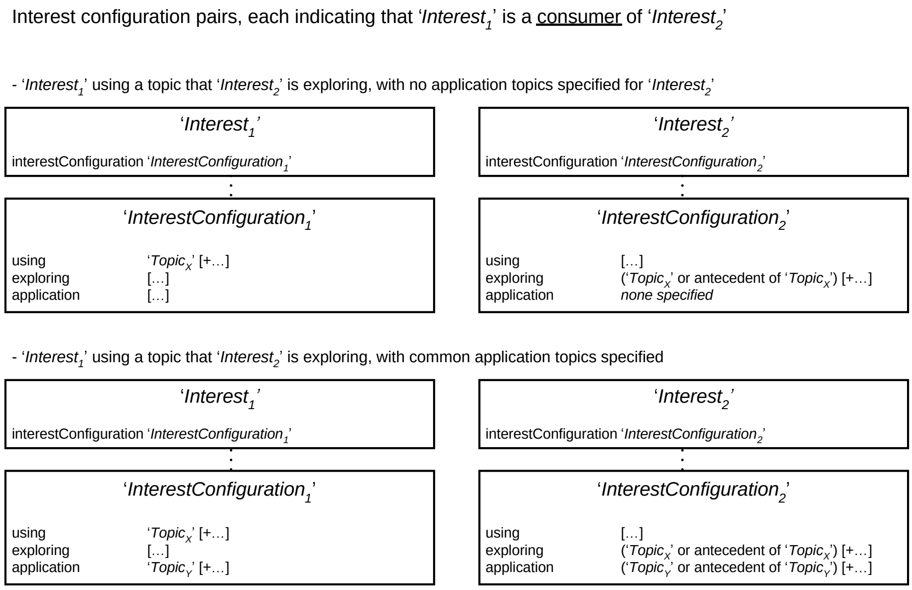


Figure 5.4 – Interest configuration pairs indicating a consumer relationship

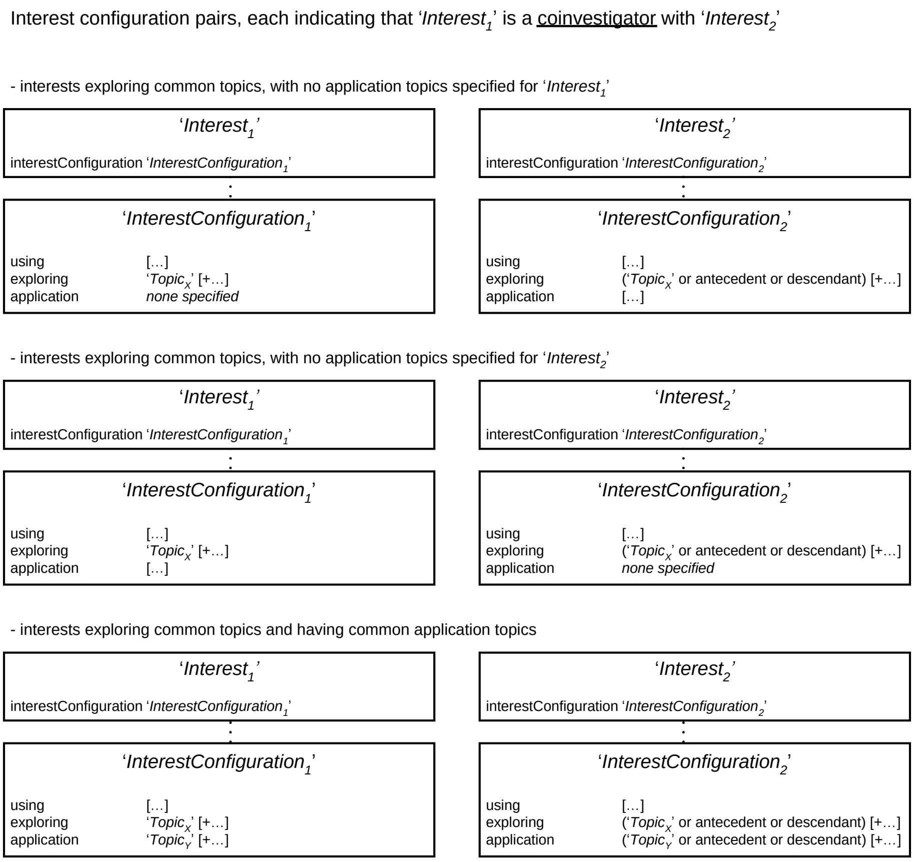


Figure 5.5 – Interest configuration pairs indicating a coinvestigator relationship

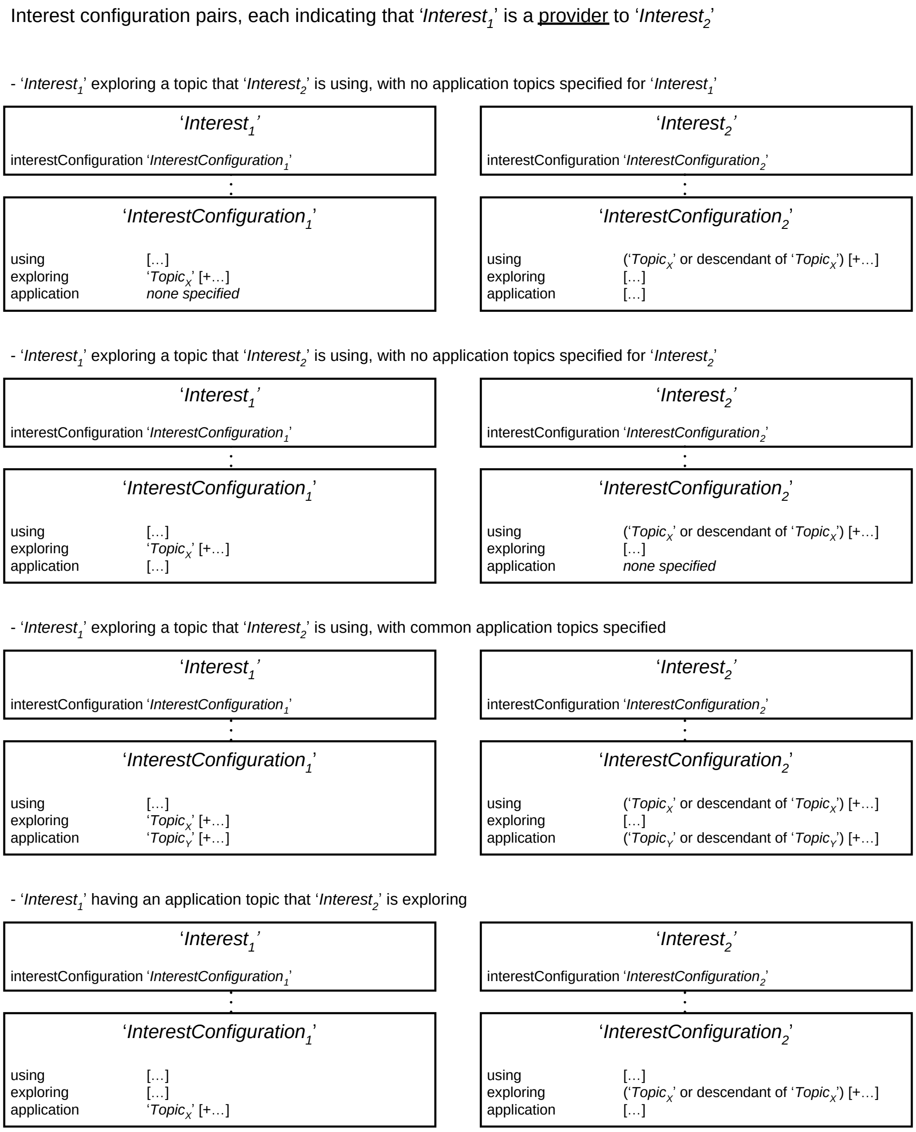


Figure 5.6 – Interest configuration pairs indicating a provider relationship

Interest configuration pair evaluation example:

Evaluation of the extent that Tom's Interest, '*Sensor Nets*', is a consumer of a colleague's Interest, '*Low-Power Wireless Deployment*'

'*Sensor_Nets'*

interestConfiguration '*SensorNetsConfig*$_1$'

:

'*SensorNetsConfig*$_1$'

using 'Bluetooth'
exploring 'Sensor_Networks'
application

'*Low-Power_Wireless_Deployment'*

interestConfiguration '*LowPowerWirelessDeploymentConfig*$_1$'

:

'*LowPowerWirelessDeploymentConfig*$_1$'

using
exploring '*Bluetooth*', '*802.11b*'
application

Matching consumer pair template, reproduced from Figure 5.4

'*Interest*$_1$'

interestConfiguration '*InterestConfiguration*$_1$'

:

'*InterestConfiguration*$_1$'

using '*Topic*$_X$' [+…]
exploring […]
application […]

'*Interest*$_2$'

interestConfiguration '*InterestConfiguration*$_2$'

:

'*InterestConfiguration*$_2$'

using […]
exploring ('*Topic*$_X$' or antecedent of '*Topic*$_X$') [+…]
application *none specified*

Topic-pair relationships to consider:

'Bluetooth' to 'Bluetooth': Topic-pair familial relationship strength = 1.0
'Bluetooth' to '802.11b': Topic-pair familial relationship strength = 0.0

1.0 > 0.0, therefore, taking strongest Topic-pair relationship, consumer relevance of 'Sensor Nets' to '*Low-Power_Wireless_Deployment'* = 1.0

Figure 5.7 – Example interest configuration pairs consumer relationship evaluation

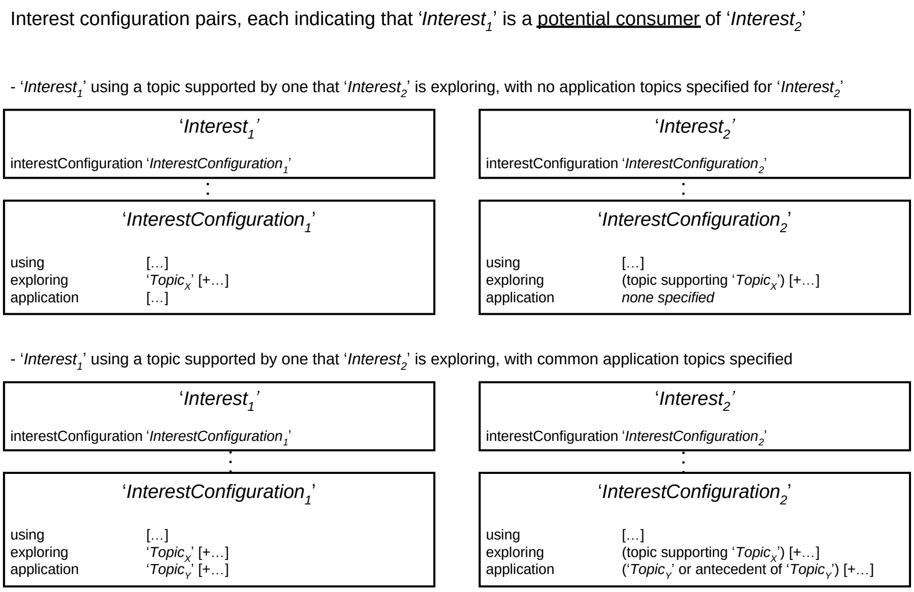


Figure 5.8 – Interest configuration pairs indicating a potential consumer relationship

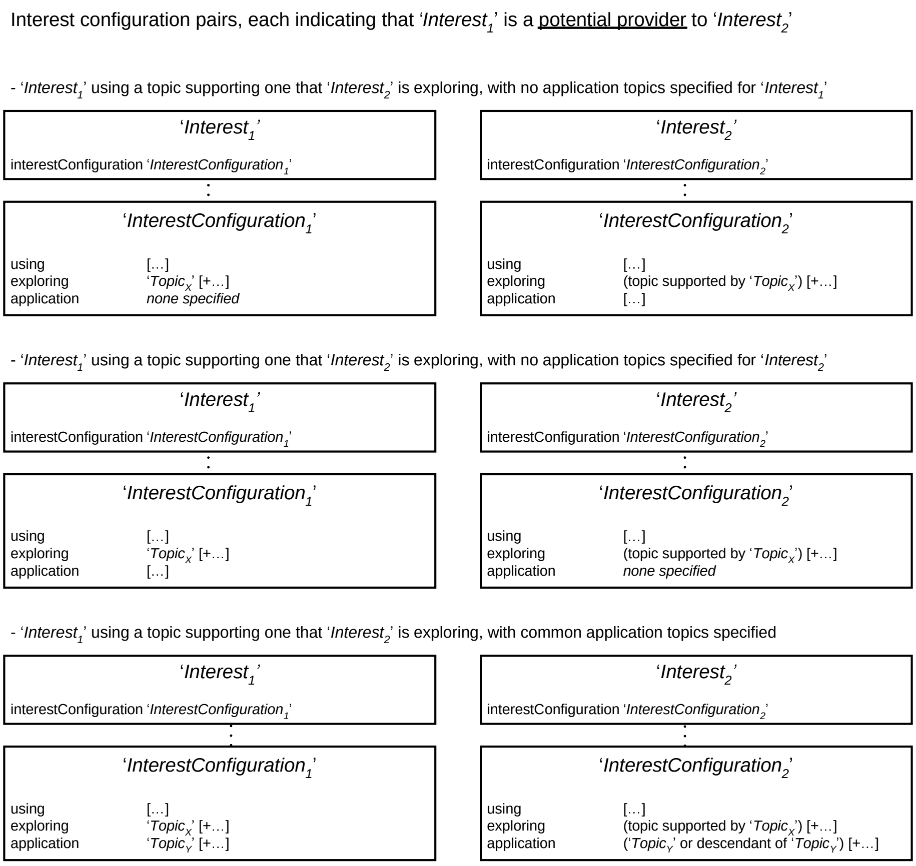


Figure 5.9 – Interest configuration pairs indicating a potential provider relationship

The core relationships used to identify Interest pairs are based on explicit declarations of Topic usage or application within that pair of Interest profiles. As shown in Figures 5.8 and 5.9, there are several additional Interest slot relationships that suggest *potential* provider and consumer pairings. These are based on the system-inferred, support Topic relationships – they suggest Interests that might be relevant to one another because of the way that the Topics they are exploring are linked in other Interest profiles. Whether these potential relationships should be considered depends on the purpose for which an Interest's pairs are being used. It would be appropriate to speculatively include any individuals who had declared an Interest in '*Mobile Computing*' when advertising an event for '*Bluetooth*' users, regardless of whether they had explicitly stated that they were using this technology. On the other hand, if the aim were to identify a Community of Expertise, it would be desirable for the system to discriminate in favour of Interest profiles that explicitly declare a relationship with the Interest in question. If the user is looking for a '*Bluetooth*' expert, they would be expected to prefer an individual who had declared a specific Interest in that Topic over one who had only stated a general Interest in '*Mobile Computing*'.

Accordingly, the 'potential strength' values are not reflected in the *providerStrength* and *consumerStrength* *InterestPairScores* slots. Instead, the *InterestPairScores* facts are augmented with two additional value slots, *potentialProviderStrength* and *potentialConsumerStrength*. These slots are populated using the same InterestConfiguration pair evaluation process as the standard consumer and provider strength slots, but the potential consumer evaluation considers both the configuration pair templates of Figure 5.4 *and* those of Figure 5.8; similarly, the potential provider evaluation considers the templates shown in Figure 5.6 *and* those in Figure 5.9.

5.2.3 Interest Group Pair Relevance Scoring

While `Interest` pair scoring allows users to target or identify other individual users, `InterestGroup` pair scoring provides group-wide generalizations that indicate the relevance that one group's current collective activities have to the collective membership of another. As with the *`InterestPairScores`*, the aim is not to accurately quantify the contribution that one `InterestGroup`'s work makes to another; the evaluations merely represent an estimate of how interesting a member of one group might find the work of another. As described in the previous chapter, although a human-authored, plain-text title and description are attached to each `InterestGroup` for users' convenience, the system evaluates them as collections of member-contributed `Interests`, that are bundled and associated with the group as `ContributorInvolvement` facts. Therefore, in determining the relationship of one `InterestGroup` to another, it is these contributed member `Interests` that are considered (Figure 5.10).

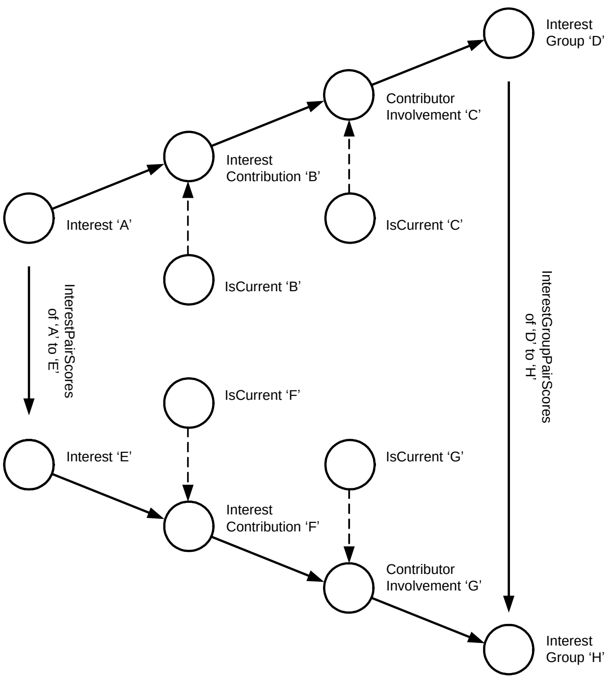


Figure 5.10 – InterestGroup pair identification

Naturally, as *InterestGroupPairScores* are based on *InterestPairScores*, the same primary relationship axes are evaluated: the extent to which an InterestGroup might provide support to another, is conducting work that parallels that in another and might be exploiting work carried out in another. They are maintained in the *providerStrength*, *coinvestigatorStrength* and *consumerStrength* slots of the *InterestGroupPairScores* fact. Users are, of course, free to contribute the same Interest to more than one InterestGroup simultaneously – for pair evaluation this special case is treated as a pair of Interests coupled by a coinvestigator relevance of 100%.

Figure 5.10 illustrates a relationship suggested where there is a link between sole Interests contributed to each InterestGroup. In practice, an InterestGroup will usually comprise several individuals, each contributing one or more of their Interests. Some Interests in one InterestGroup might have *InterestPairScores* links to several Interests in the other InterestGroup, while others might be 'orphans', that are of no relevance whatsoever to this second InterestGroup. As shown in Figure 5.11, these links are scanned to determine the strongest overall inter-group relationship for each of the two InterestGroups' members and an average of these values calculated to give an overall relevance score for the pair InterestGroup. Let us consider the implications of three possible combinations of resultant pair InterestGroup score:

- *low scores for both pair* InterestGroup*s*: if there are any strong inter-group relationships, they relate only to a small proportion of members in each group. While these links may be important to the individual group members concerned, it cannot be assumed that the group memberships as a whole will be interested in each others' work.
- *both pair* InterestGroup*s scored highly*: each of the groups has broad relevance to the other - it is likely that that membership of each group would wish to be made aware of the existence and activities of the other.

- *one* `InterestGroup` *of the pair scored low, the other scored highly*: one group is focussed on work that is relevant to a part of the other group's activities.

In this third case, it is not immediately obvious where the balance of interest would lie between the two groups. There are likely to be fewer `InterestGroup`s dedicated to a particular narrow field than there are ones that include this field among a wider range of `Interest`s; this might suggest the higher scored group would be more relevant to the lower scored one than vice versa. However, the area of overlap or mutual interest between the groups directly concerns a larger proportion of members in the highly scored group than in the lower scored one. This is illustrated in Figure 5.12 – it is clear that, while, for instance, a message summarising group A's activities would be relevant to all

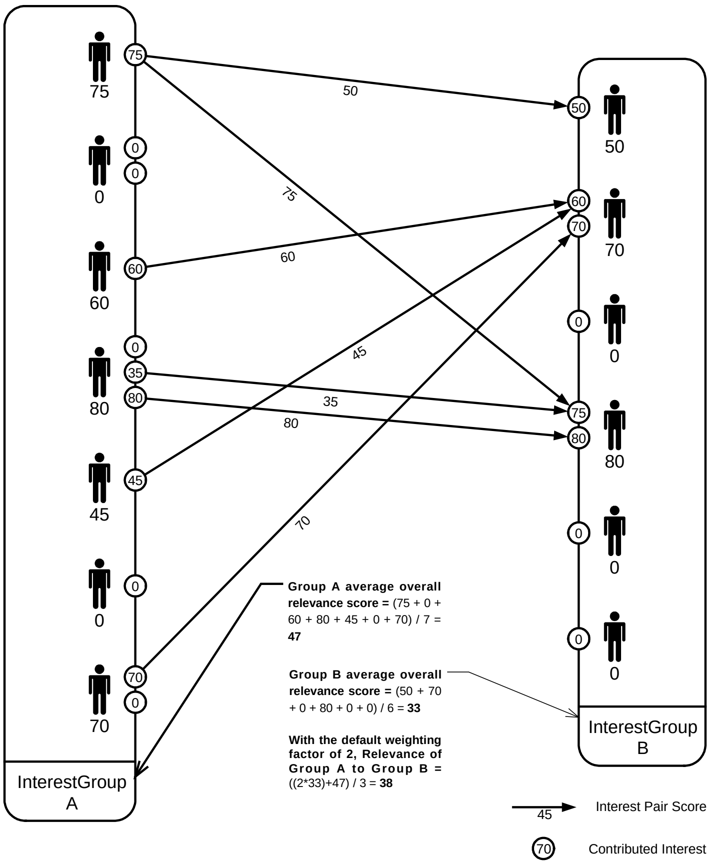


Figure 5.11 - InterestGroup pair scoring

members of group B, one describing the work of group C could only be assumed to interest about half the membership of group D[1]. The system accounts for this by weighting the overall group pair score in favour of the second, 'to' column score, as shown in Figure 5.10. By default, a static weighting factor of 2.0 is used. This factor can be freely adjusted by the system administrator to increase or reduce the awareness of niche neighbour groups that will be afforded to broadly based groups. A future enhancement of the platform could allow automatic regulation of the weighting, perhaps using a reduced value to evaluate relationships between recently formed `InterestGroup`s.

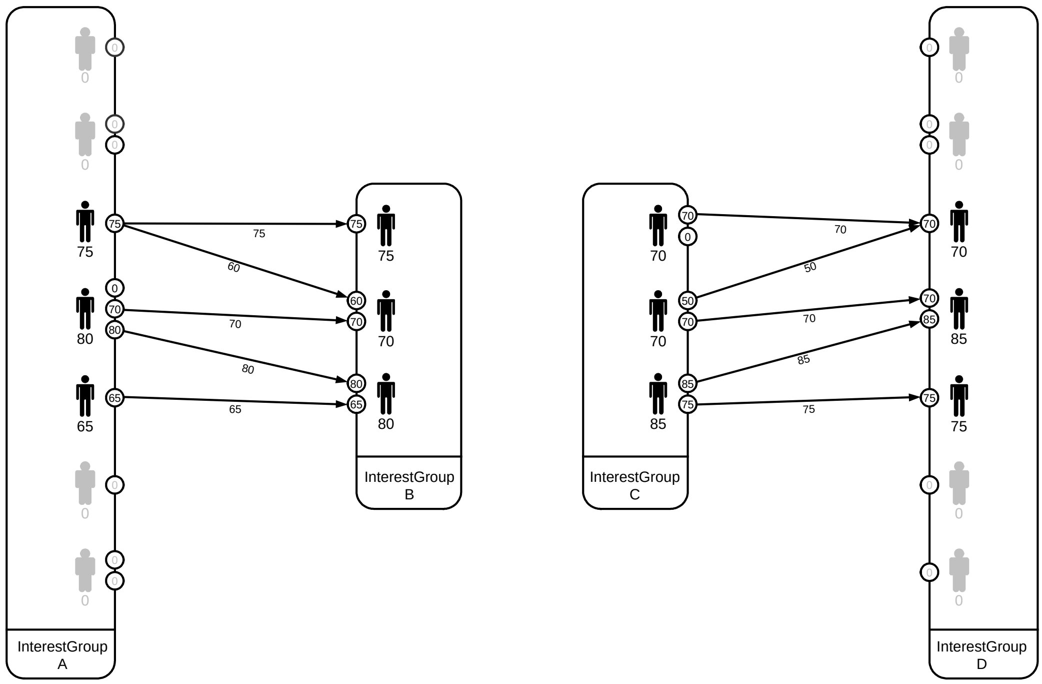


Figure 5.12 - InterestGroup overlap examples

[1] Though group pair scoring is predicated on the *`InterestPairScores`*, it does not replace them -- both are exposed and in many instances, such as this, it would be more appropriate for the user to target specific involvement relationships using the *`InterestPairScores`*, rather than rely on weak group-wide similarities.

5.2.4 A Note on Scalability

Care has been taken to ensure that the network of automatically-maintained relationships remains manageable as the system user-base grows. The use of separate 'checked' and 'modified' timeslots in the pair-score facts ensures that any modification of the fact-base does not precipitate unnecessary re-evaluation of unaffected relationships. Pair-score facts are augmented with utility slots that register identifiers of 'keystone' `Topic`s, *`TopicPairScores`* or *`InterestPairScores`* facts upon which the pair evaluation depends; therefore when one of these facts is retracted, only pair-scores that depended on that fact must be re-evaluated. The number of relationships that are maintained is minimised as pair-score facts are stored only for non-zero pair relationships – existing pair-scores are retracted if modification of the fact-base eliminates the relationship to which they refer.

5.2.5 Target Profile Population

`TargetProfile`s allow a search query, for users that meet a particular set of criteria, to be lodged with the system, causing a dynamically updated set of results to be maintained in the fact-base for other inference processes to use. `TargetProfile` maintenance involves the identification of sets of `Person`s, current `Interest`s and `Involvement`s and `DiaryEntry`s that meet the selection criteria laid out in the profile itself and in any attached `InterestTarget`s and `GroupTarget`s. As population of the profile requires attached `InterestTarget`s and `GroupTarget`s to be up-to-date, let us first examine the process by which these are maintained.

`InterestTarget`s and `InterestGroupTarget`s provide lists of current `Interest`s or `Involvement`s that have a significant relationship with the root `Interest` or `InterestGroup` identified in the target. The pair-score facts maintained in the fact-base record every non-zero pair-wise `Topic`, Interest and InterestGroup relationship; inevitably many of these comparatively weak pair-scores will be irrelevant to users, but if they were simply removed from the fact base, the rules that detect a potential pair relationship would force them to be recreated, placing the system in an infinite loop[2]. Therefore, so that significant Interest and InterestGroup pairs can be identified quickly for InterestTarget and InterestGroupTarget population, the Interest and InterestGroup facts hold thresholded pair sets, which are recompiled as necessary, whenever the pool of relevant pairs from which they are drawn is modified. InterestTargets are populated simply by tracking the contents of those thresholded pair slots in the root Interest indicated by their match selection flags (matchCoinvestigators, matchConsumers and matchProviders). InterestGroupTargets are populated in two stages. Firstly, a list of matching InterestGroups is obtained by cloning the relevant thresholded pair slots in the root InterestGroup; if the InterestGroupTarget matchGroup flag is TRUE, then the root group identifier is also added. This list is then tracked to maintain a second slot in the InterestGroupTarget referencing the Involvements of all the matching groups.

To populate the TargetProfile, the Interest and Involvement matches indicated by the attached interest- and group- targets are assembled, together with DiaryEntrys that pertain to any SocialGatherings referenced in the profile itself, to give lists of Interests, Involvements and DiaryEntrys that satisfy *an* aspect of the profile specification. The Persons with which the listed items are associated are identified; the system can then identify Persons that satisfy *all* aspects of the profile and any Interests, Involvements or DiaryEntrys belonging to those Persons form the final list of profile matches. Repopulation of the TargetProfile is triggered by direct

[2] Although, theoretically, this could be addressed with more complex trigger rules, this would be unworkable in practice because, as complex mathematical evaluation cannot be optimised by the inference engine, modification of a single Topic, Interest or InterestGroup would force re-evaluation of every pair-wise relationship.

modification of the profile itself, repopulation of an attached GroupTarget or InterestTarget, or a modification of a DiaryEntry that reflects a change in attendees of listed SocialGatherings.

The temporal component of the TargetProfile is gained by virtue of the TargetProfile being a subclass of MultiTemporalThing. The attached TimeIntervals determine the time period during which the TargetProfile is to be considered valid; when active, its *current* slot is set TRUE. When this flag is unset, the TargetProfile is no longer considered to apply and any Bulletin posts and event or informal group invites that are predicated on it are revoked.

5.2.6 Bulletin Management

Bulletins are addressed using their attached TargetProfiles, so the initial posting decisions are performed using the TargetProfile population mechanism described above[3]. By itself, though, the TargetProfile only provides a snapshot of all Interests, Involvements, DiaryEntrys and Roles that currently satisfy the bulletin's addressing criteria – as the system constantly assesses possible TargetProfile matches, it cannot be assumed that any of these matches necessarily occurred at the time the Bulletin was posted. Therefore, so that a user's Bulletins can be displayed in the order that they became visible to *them*, the matches indicated by the TargetProfiles are used to populate *bulletinBoard* slots in current Interest, Involvement, DiaryEntry and Role facts. Bulletin identifiers are appended to a *bulletinBoard* slot's contents, immediately the corresponding fact's identifier

[3] MediaBulletin postings may be made to individuals who will not be permitted to view the MediaItem that is being promoted. Ordinarily a client application will be expected to disregard the MediaBulletin if it determines that its user cannot access the MediaItem, though in a meeting situation where items are tabled for discussion, it may instead be appropriate to show the item 'padlocked', so that the user can choose to negotiate access with its owner.

is listed as a match in the `TargetProfile`, therefore the board structure reflects the order in which `Bulletin`s are posted to it.

As a populated *bulletinBoard* slot contains a list of pointers to original `Bulletin`s, not local copies of the `Bulletin`s themselves, a `Bulletin`'s owner (*its author, if it is a personal bulletin, or, in the case of a group bulletin, any member of that group*) is free to make any changes they wish to its contents; these will be visible to anyone who then looks at the `Bulletin`, regardless of whether they had already viewed it before it was altered. Previous versions of the `Bulletin` are not retained, but the identities both of its original creator and the author of the most recent version are recorded. A `Bulletin` can also be retargeted by its owner at will, by modifying an existing attached `TargetProfile` or defining a new one.

Further to the interest, involvement and event-specific bulletin boards, each `Person` fact also has a *bulletinBoard* slot that carries all messages visible to that individual, allowing them all be retrieved as a single list, temporally sorted in order of first display. This function will typically be used by a client in conjunction with a list of previously read `Bulletin` identifiers, to obtain an ordered compilation of all 'new' `Bulletin`s.

Just as out-of-date notices are removed from traditional physical bulletin boards and discarded, the system deletes `Bulletin`s when their associated `TargetProfile` expires (*i.e., when the* `Bulletin`*'s historic slot is set TRUE*) and removes any references to them in the fact-base. Alternatively, a `Bulletin` owner can opt to delete the message manually before its display period expires; again any trace of the message will be removed from the system.

Though the primary function of the bulletin board sub-system is to post user-authored bulletins, it also automatically generates and posts 'digital flyers' to publicise new water cooler groups and forthcoming events. InformalInterestGroup and SocialGathering facts possess an advertisement utility slot that holds the identifier of the corresponding flyer Bulletin. Flyer generation is initially triggered by a rule that detects facts whose advertisement slot is empty. The title and message (name and descriptiveText) of the flyer Bulletin are obtained from the name and first paragraph of the descriptiveText of the InformalInterestGroup or SocialGathering fact – with an event flyer, the TimeInterval of the SocialGathering is also parsed and the event time and date added to the flyer message. In the event that any of these details are changed after the initial creation of the flyer, the Bulletin is updated automatically, triggered by a rule that compares the modification dates of flyers and the objects that they advertise.

Clearly it is desirable that flyers are only distributed to individuals who are invited to join the water cooler group or event in question. This is achieved by deriving their TargetProfiles from those referenced in the InformalInterestGroup or SocialGathering advertiseTo slot. The flyer's TargetProfile is associated with a RelativeTimeInterval related to the InformalInterestGroup or SocialGathering to which it refers; a water cooler flyer is timed for display during the lifetime of the InformalInterestGroup, while an event flyer will be displayed for the period leading to the SocialGathering. If a flyer recipient joins the InformalInterestGroup or confirms they will attend the SocialGathering to which the flyer refers, the Bulletin identifier is added to that Person's *concealedMessages* slot, to indicate to client applications that that particular bulletin is superfluous and need not be shown.

5.3 Client-Initiated Functions

In addition to the rule-based, data-driven operators described above, a package of client-initiated functions is included within the core server functionality. Primarily, these allow client applications to retrieve properties of objects held within the conceptualisation and to manipulate the conceptualisation, to create new objects or modify the properties of those that already exist. Four bundles of manipulation meta-functions are provided, simplifying creation and management of bulletins, social gatherings, interest groups and personal profiles (*posting a new bulletin, for example, requires the creation of the* `Bulletin` *object itself, a* `TargetProfile` *and associated* `TimeInterval` *and, typically, some combination of* `InterestTargets` *or* `InterestGroupTargets` *– the 'bulletin creation' function creates and populates all these objects and establishes the necessary references between them*). Each function is guarded by an authentification barrier that ensures that clients are only able to access of modify objects of which they have ownership rights. More complex client-initiated functions provide applications with straightforward means to identify *expertise* and *social networks*. The operation of these functions is described in the following sections.

5.3.1 Expertise Identification

The Expertise Identification function returns sets of core and boundary interest and likely expertise for a given root `Interest` profile or `Topic`. `Interest`-rooted expert sets are identified using the `InterestPairScores` facts that are maintained for target profile population. To allow the compilation of sets of individuals that have interests relevant to a single `Topic`, however, a set of `Topic`-to-`Interest` pair scores (`TopicInterestPairScores`) must also be obtained. Although these scores are not required by the data-driven server components discussed previously, they, too, are maintained dynamically; from the emphasis placed on Communities of Practice in workplace evaluation literature it was judged that the `Topic`-root expert function would be sufficiently well-used to justify the additional server overhead incurred. For the purpose of pair score calculation, a `Topic` is considered equivalent to an `Interest` *exploring* the `Topic`; the scoring algorithms used are a subset of those employed to score `InterestPairScores`.

A stored query is executed to retrieve a list of current pair scores of the root Interest or Topic and hence a list of individuals' current Interests to which the root is relevant. A composite relevance score is obtained for each Interest from the Pair supports, supportedBy and coinvestigator scores, weighted in favour of the strongest relationship.

Each Interest is *typed*. In accordance with the conclusions of Lindgren et al. [Lindgren03], those that have a composite score > 75% *and* have been contributed to a formal group in the past 18 months are regarded as potential evidence of *expertise*. If the candidate Interest has been reconfigured since the involvement (i.e a new InterestConfiguration has been attached), coinvestigator pair strengths of the root Interest or Topic to the candidate as it was configured during the period of involvement are evaluated. Then, if the coinvestigator pair strength of the root to the candidate Interest for at least part of the involvement is also > 75%, the Interest is confirmed as signifying likely *expertise*.

Non-expert Interests with an overall current direct relevance > 50% are typed, simply, as *interest*. The remainder are classed as *possible interest*. They are then sorted into person groups – an individual may, of course, have more than one relevant Interest – with the strongest Interest type being passed to the person; i.e. an individual who has one relevant Interest typed *expertise* and another typed *possible interest* will be considered *expert*. Finally, to avoid an excessively large number of 'possibly interested' individuals being returned, a threshold is applied. Any Interest with overall relevance below the median average and below 50% is discarded. In line with Lindgren's research, thresholding does not favour expertise over interest.

Previously we have identified the importance in innovation-led work of the 'boundary member' – an individual who may be on the periphery of a particular field, but is better able to relate the subject to one outwith that field because other shared interests give them common ground for discussion; "*people are unlikely to share some new topic… unless they share some common ground*" [Carter03]. Therefore, in addition to the core expertise set, for each of the user's declared `Interest`s a boundary set is also identified.

A boundary set is a subset of the unthresholded core expertise set, containing only those individuals that also have one or more `Interest`s relevant to the user's `Interest` on which it is based. Against each core set member a stored query is executed to retrieve any *`InterestPairScores`* facts *of* one of their `Interest`s *to* the user `Interest` in question. Using the formula above, an overall relevance is calculated for each member `Interest` referenced. Individuals with no links to the user `Interest` are then discarded from the boundary set.

Once again thresholding is required to ensure that the set of individuals returned is not meaninglessly large, but here the aim is to include those individuals that have strong or moderate interest in either the user's `Interest` *or* the root `Topic` or `Interest`, while eliminating those with comparatively little interest in both. Inevitably, then, a more complex thresholding regime must be applied:

1. Two threshold values are determined:
    a) The *root relevance threshold* represents either the median value of each individual's maximum root relevance score or, if the smaller value, 50%.
    b) The *boundary relevance threshold* represents either the median value of each individual's maximum boundary relevance score, or, again, if the smaller value, 50%.
2. We disregard any individual that whose maximum root-relevant `Interest` is scored below the root relevance threshold *and* whose maximum boundary-interest-relevant `Interest` is scored below the boundary relevance threshold.

3. For each remaining individual:

   a) If they have one or more root-relevant `Interest`s scored above the root relevance threshold, exclude all those that are scored below this value. If not, dereference all but the most relevant `Interest`.

   b) If they have one or more boundary-interest-relevant `Interest`s scored above the boundary relevance threshold, exclude all those that are scored below this value. If not, dereference all but the most relevant `Interest`.

Finally, each expertise set identified is sorted, firstly into type groups and then by member family and given names. The relevant `Interest` identifiers are included in the returned sets package together with those of the member `Person`s, enabling the client application to present to the user the rationale behind each member's inclusion.

5.3.2 Social Network Identification

The Social Network function returns interpersonal chains that most strongly link the user to a given `Person` or `Group`. This information can be used to initiate a 'chain of recommendation', whereby the user's interest in the target `Person` or `Group` is relayed from `Person` to `Person` along the chain, or as a kind of 'letter of recommendation' that the user can use to introduce themselves directly to the target, as described by Kautz [Kautz97]. The function identifies chains of maximum length 4 nodes, or 3 degrees of separation. Some limit is necessary to ensure that only socially meaningful chains are returned. Travers' and Milgram's [Travers69] results suggest a maximum limit, showing that by exploiting individual knowledge of professional contacts, a chain of 4 or 5 degrees of separation was sufficient to link any two individuals. They do not claim, though, that chains of such length are of practical use. Therefore, in line with the stated aim of the system to echo commonsense human rules of thumb, 3

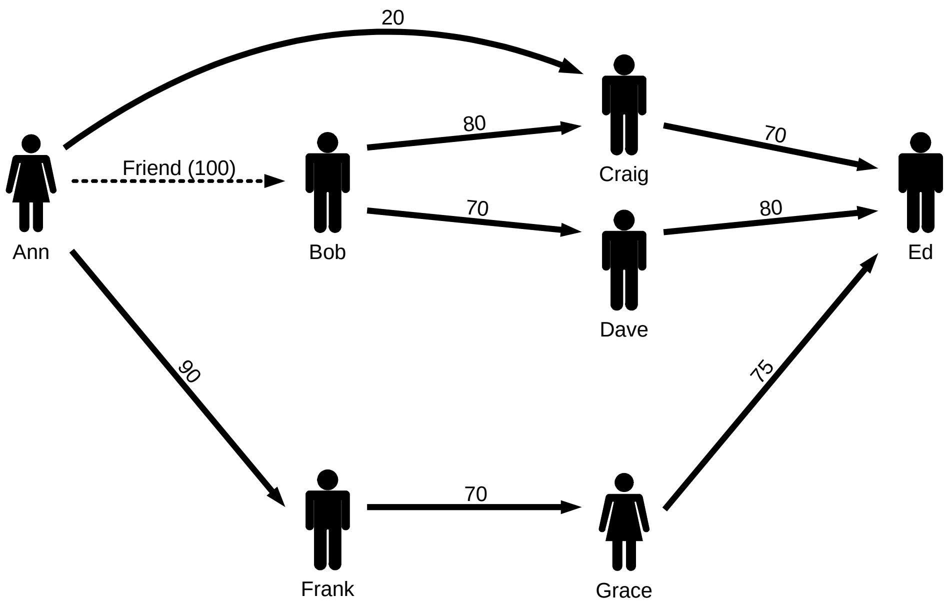


Figure 5.13 – Social network evaluation

degrees of separation was chosen as representing the largest chain that can comfortably be expressed in plain English ("*…your friend Tom is currently working with her boss, Steve…*").

Although the usefulness of social networks in innovation is clear, the likelihood that any one user will wish to identify a social network to any one other is obviously small. Given the large number of social connections that will exist within a typical user population, it does not make sense to dynamically maintain social links and chains as we do for other relationships used by the system. Instead, they are identified and evaluated as required, using a low-priority snapshot clone of the fact-base so that the processing involved does not interrupt other users' interaction with the system. Nonetheless, the dynamic pair scores that *are* maintained are available in the snapshot for use by the social link evaluation routines, so the performance benefits of a largely dynamic system are not lost to this function.

Social chains are identified using *backward chaining*, a process whereby the inference engine actively seeks to satisfy a goal condition indicated by an executed stored query – in this case, a social link from the user to the target. For chain identification, a link is considered to exist from one individual to another if they have temporally overlapping periods of membership in at least one Group, or the second individual lists the first as a friend. Let us examine the mechanism by which the chain Ann-Frank-Grace-Ed, represented in Figure 5.13, is identified (for clarity, names rather than identifiers are shown). A query is executed to identify `Connection` facts *from* Ann *to* Ed – i.e. those facts that match *(Connection (from Ann) (to Ed))*. Because the `Connection` fact template is enabled for backward chaining, the inference engine asserts a fact to represent the goal condition: *(need-Connection (from Ann) (to Ed) (via "")).* Two rules are implemented that trigger on need-Connection facts:

> If *(need-Connection (from Person_V) (to Person_W) (via $Via))* and a direct social link exists between them, and *(need-Connection (from Person_U) (to Person_W) (via ""))*, assert *(Connection (from U) (to W) (via $Via))*. Thus a match for the query is identified.

> If *(need-Connection (from Person_X) (to Person_Z) (via $Via))* and a direct social link exists from Person X to Person *Y*, and $Via currently references fewer than 3 intermediate nodes, assert *(need-Connection (from Y) (to Z) (via $Via+Person_Y))*.

Hence, each final Connection fact, returned by the query, contains the chain of Persons that comprised the connection in its *via* slot – e.g. *(Connection (from Ann) (to Ed) (via Frank Grace))*. The process repeats automatically until all chains, of 3 degrees of separation or less, from the user to the specified Person have been identified; the query then returns a vector of these chains. If the target that the user wishes to connect with is a Group rather than an individual Person, chains are sought for each of the Group's current membership in turn.

In accordance with Granovetter's definition of link strength [Granovetter73], the following factors are assessed to evaluate individual interpersonal links:

- recency, duration and timescale of collaboration
- number of collaborations
- number of instances of joint attendance at physical project meetings
- number of instances of co-authorship of media items

These are codified as sets of JESS fuzzy sets [Orchard01], representing the strength of social link that a particular value for any of these factors indicates. Each of these sets is administrator-modifiable, allowing the social network engine to be configured for the particular conditions and attitudes of the organisation in which the platform will be deployed. In the default configuration, for example, a period of collaboration of at least 2 years indicates a 'very strong' social link; in an environment characterised by very low employee churn, though, this might be considered only to suggest a 'moderately strong' connection.

Regardless of these factors, a full strength link is considered to exist from one individual to another if the second `Person` has listed the first in their `friends` slot. The link score is recorded, together with the identifiers of `Group Involvement`s, `MediaItem`s and `DiaryEntry`s that form the most significant evidence of the link's existence. This allows the nature of the link to be 'explained' to the user. Link evaluations are cached until all identified chains have been scored as different chains will commonly share one or two links (in Figure 5.6, for example, two chains contain the link from Ann to Bob).

To calculate the overall strength of a chain, the system takes into account not only the strengths of link between immediately adjacent links in the chain, but also any parallel links that have been identified between non-adjacent links. Referring again to Figure 5.13, Nardi tells us that even a weak direct link from Ann to Ed could quickly be reinvigorated, but in this case there is no direct link at all. The shortest (*geodesic*) chain is via Craig, but the link from Ann to Craig is scored very low; we can imagine, for example, that it stems from historic co-membership of a large virtual community. A more promising network to exploit might be that through Frank and Grace – though it involves an additional degree of separation, each link is scored highly, indicating active, well-established working relationships between each pair of individuals. However, though Bob and Ann have not collaborated professionally, their friendship yields two additional strong connections to Ed; one through Dave, the other via Craig. This second chain is assumed to offer the strongest overall connection from Ann to Ed, as a recommendation from Bob to Craig would reinforce the weak direct social link that already exists between Craig and Ann.

As Granovetter observed, in determining the most effective social bridge to an individual or community, we must consider not only the strengths of individual social ties at the present moment, but also the number of social ties that must be traversed [Granovetter73]. We would expect a reasonably strong two-step chain to be preferable to a slightly stronger three-step one, so the overall chain scores are scaled according to the number of links that they contain. Additionally, where the user has requested networks to a Group, the chain score is weighted according to the duration and recency of the chain terminator's involvement in that Group. Again, the degree of weighting applied can be modified by the platform's administrator. Finally the chains are sorted according to their score and the five strongest returned.

5.4 Summary

Pair-wise relationships between topics, interests and interest groups form the basis of a dynamic, rule-maintained network of organisational relationships. In conjunction with rich query-driven functionality, also implemented within the reasoning engine, they allow a comprehensive suite of awareness-, community- and communication-led services to be derived from a single organisational model. In the next chapter, we will explore a client application scenario that leverages these services to satisfy the requirements for innovation support summarised in Chapter 4.

Chapter 6

A Kiosk Client Application

The Aware platform provides a dynamic organisational model and some primary functions with which its relationships can be evaluated. Few assumptions have been made thus far of the way that users will engage with the platform; this will largely be shaped by the compositions of functions presented in client applications. This chapter examines a possible deployment of the platform within a technical research organisation. Drawing on the organisational analyses and software evaluations presented in previous chapters, situations of use are identified from which a client application suite is specified. The implementation of a key component of the client suite, a kiosk interface for semi-public spaces, is then described.

6.1 Developing a Client Application Suite

In Chapter 4 we briefly considered 'situatedness', the recognition that the requirements a user has of a system vary depending on their current context. The community-based KM tools discussed in Chapter 3 addressed several situational types found within research organisations. Ontocopi supported private, *desk-based* research. NuggetMine also assisted desk-based work, notably as a web browsing companion. Drehscheibe, GroupCast and hebb targeted semi-public spaces, supporting informal interaction and awareness through fixed open-access *kiosk displays*. ProxyLady and Proem placed lightweight messaging and interpersonal awareness services on a personal wearable device; they invited *mobile* use. Though not represented in the tools surveyed, the importance of face-to-face discussions necessitates consideration of a fourth situation, the *meeting room*[1]. Broadly, the activities performed in these situations can be characterised as follows:

- Desktop – individual work activity
- Kiosk – individual engagement in the community
- Meeting Room – group work activity and engagement in the community
- Mobile – immediate individual engagement in the community

Four corresponding client packages are suggested below, each emphasising different aspects of the platform functionality.

6.1.1 Desktop Client

The Desktop client will emphasise support for personal information management and tasks that are embedded within the user's everyday work. It will allow the user to create and manage their personal interest profiles, assembled by selecting topics from the taxonomies held in the system. It will also allow media items to be browsed and facilitate exchange of URL Bulletins, by way of a hook into the user's web browser client. Exploiting the platform's community of interest function, the user will also be able to locate individuals knowledgeable about their current task.

6.1.2 Kiosk Client

The Kiosk client will offer the user a point of engagement with the wider organisation, supporting electronic analogies of some of the physical social tasks that are carried out in the workplace. It will be a space for public information

[1] Numerous applications have been developed to enable replay of meeting room proceedings after the event or to facilitate meetings between different physical locations. Here, though, we are concerned with activity performed *within* a meeting space.

exchange, where bulletins can be posted and browsed and informal watercooler groups formed, identified and joined. It will also provide a window through which to 'browse' the organisation, to gain awareness of the activities of groups and individuals.

6.1.3 Meeting Room Client

The Meeting Room client will augment the functionality provided by the public kiosk client with tools that enable a group to engage with the organisation. Specifically, it will allow the creation and management of a formal group and of group events.

6.1.4 Mobile Client

Exploiting the platform's ability to push notifications to client applications, the Mobile client will provide a lightweight interface with timely or time-critical information, such as last-minute changes to an event that the user will be attending. It will also be able to alert the user to physically proximate individuals that have newly acquired related interests.

6.2 Kiosk Client Implementation

The construction of a complete suite of application tools is beyond the scope of this thesis. Indeed, in a practical deployment, it is likely that some of the tools would also package access to other organisational systems, such as Instant Messaging or employee position location, so it would be unrealistic to present the work as a turn-key solution. The construction of a functional prototype of one of the suggested tools, however, *is* a worthwhile exercise as a component of the platform evaluation. It can place the services provided by the platform in a realistic operational context, demonstrate how services can be combined to provide a holistic organisational awareness experience and illustrate that the complex organisational model maintained can be accessed in a simple and intuitive fashion. This section describes the implementation of a prototype kiosk interface for semi-public spaces.

6.2.1 Kiosk Design Considerations

The design of a kiosk-based client requires an understanding of the interaction modalities that a kiosk can support. Though kiosks are immediately recognisable by one element – the provision of a large, fixed, conventional display that invites interaction – they vary considerably in their technical configuration. A common basic kiosk arrangement is a single, standalone, touch-operated display. This has the advantage that it requires a minimum of space and can be mounted entirely behind glass, allowing interactive displays to be provided in public situations such as shop windows and street-side information points. At the other extreme, a deployment might consist of row upon row of standard desktop PCs, distinguished from a conventional office space only by the provision of some form of token-based access mechanism. This arrangement is exemplified by the giant Internet cafes operated by the easyInternetcafé[2] chain.

Recently, two technologies have been launched to combine the richer interaction offered by traditional keyboard and mouse with the operational convenience of the standalone screen. Virtual keyboard[3] devices can superimpose an optically-sensed keyboard onto any flat surface. An infra red laser creates a 'detection plane' of infra red light immediately above the surface. A CMOS camera sensor and custom image processing hardware are used to map the detection plane. When an object – a finger – passes through the plane to touch the physical surface below, its position is determined, compared against reference points in the field that comprise the virtual key areas and a corresponding keystroke generated. To orient the user, a separate visible laser beam can be used

[2] www.easyinternetcafe.com

[3] www.vkb-tech.com

to project a Qwerty keyboard outline onto the table surface. Alternatively, if the virtual keyboard is fixed relative to the surface, the keyboard outline can simply be marked out permanently on the table.

Virtual keyboards are ideal for kiosks situated in cafeteria or coffee-shop environments. The absence of conventional keyboard maximises the table area available for eating, while if a particular kiosk interaction requires keystroke input, the user can clear their food to one side and use the table surface as a keyboard. The other interaction technology developed can be used where there is no convenient flat physical surface available to double as a keyboard. It instead exploits a ubiquitous piece of personal hardware as an interaction device:

> "*Using High Energy Magic's SpotCode Platform you can use your mobile* [camera] *phone to easily interact with computers of all shapes and sizes: whatever you're doing, wherever you are. Your phone becomes your own all-in-one mouse, keyboard, storage device and authentication system*"[4]

The SpotCode is a circular barcode that can be displayed as a component of a kiosk user interface element and read by a camera phone on which SpotCode reader software has been installed. When the user points their phone at the tag, it is identified by the phone and a partnership established between the phone and the kiosk platform via a Bluetooth wireless connection. The user can then interact with the tagged user interface using their phone's keypad and touch screen. Interface elements can also be "*thrown*" from the kiosk to the phone display, allowing private interactions with the publicly viewable kiosk application.

These novel interaction technologies make the kiosk a highly flexible hardware platform, still able to offer a lightweight, immediately available access point into an organisational infrastructure, but also capable of effectively and

[4] www.highenergymagic.com

seamlessly supporting more complex tasks when required. An application intended initially for a 'public-use PC' kiosk will be transferable to wall-mounted interactive displays. The availability of some form of keyboard entry and pointer device can be assumed. However, if a client is to accommodate different kiosk types, it must support different screen sizes. As modern display technologies have a native resolution that varies depending on the size of the screen, this implies that a vector-based user interface implementation will be most appropriate.

6.2.2 Kiosk Client Architecture

The most widespread vector-based technology for thin client construction is Macromedia Flash[5]. Conveniently, Flash also has good support for XML messaging over TCP/IP sockets, so can directly interface with the server platform without the need for a proxy. Accordingly, Flash was used to implement the kiosk client application. To ensure that the user interface did not unduly constrain the paths that users could take between the various platform functions exposed, a window-based interface was designed. To support this a custom window manager has been implemented from graphics primitives. Figure 6.1 is an overview screenshot that shows the client in operation. The implementation of the individual function windows and other supporting aspects of the client interface are discussed in the following sections. What follows is not presented as a definitive solution, but as a suggestion of how platform functions can be leveraged to construct a plausible client application.

[5] www.macromedia.com/flash

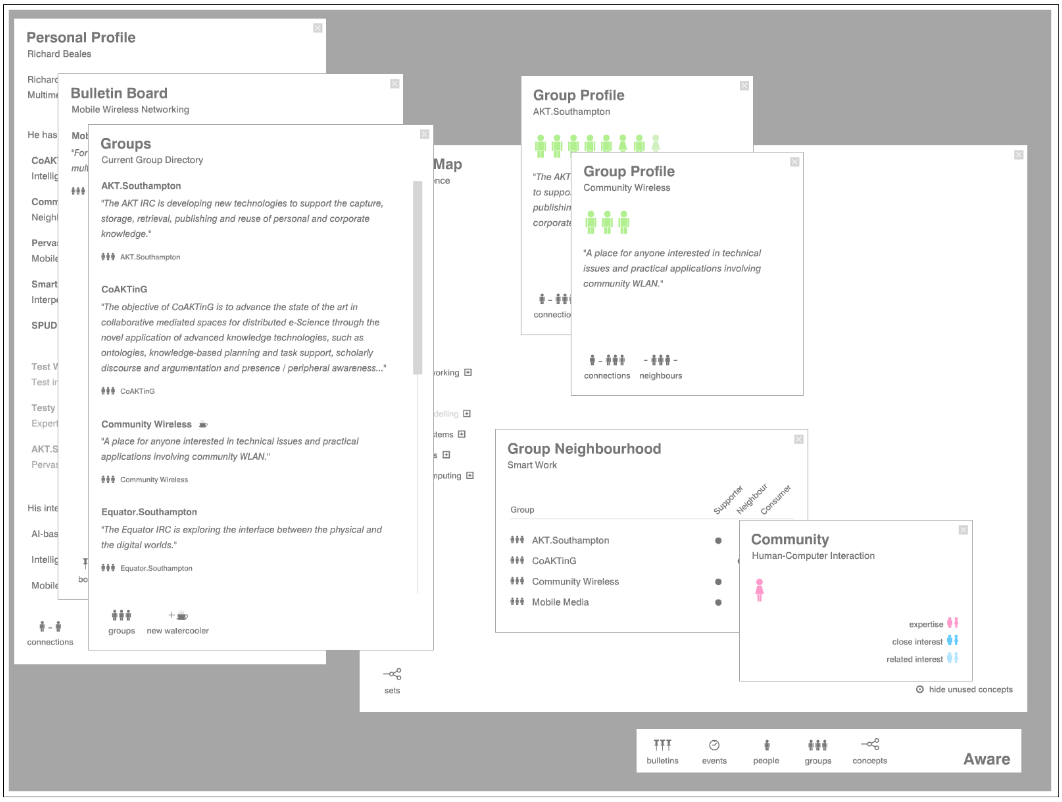


Figure 6.1 – Kiosk Client Workspace

6.2.3 User Authentification

Recognising that users will likely grab a few minutes to interact with the kiosk over a snatched coffee, instantaneous login and logout are provided, using convenient iButton[6] authentification. iButtons (Figure 6.2) are small metal canisters, typically carried on a key ring or dog tag, that hold a unique 64 bit identifier; they are widely used as a means of employee identification within organisations. A utility slot in the ontology `Person` representation has been provided to store this identifier, allowing the server to match the numeric code to an individual platform user. To login to the kiosk, the user latches their iButton into the receptacle provided; likewise, when they wish to logout, they simply unlatch their iButton. On doing so, the kiosk client serialises the user's current workspace and transmits it to the server platform, where it is stored within the

[6] www.ibutton.com

JESS fact base. When they next login via a kiosk, it is retrieved and the workspace recreated. Thus, if a user is interrupted part way through a task, they can resume exactly where they left off later. Though not enabled in the current client implementation, it would be easy to allow the user to maintain multiple workspaces, so that they could separate tasks pertaining to different activities or involvements. As they are centrally stored, the workspace serialisations can also be referenced to configure other client interfaces.

Figure 6.2 – iButton and Kiosk Authentification Screen

6.2.4 Toolbar

Access to the principal client functions is provided by a floating toolbar (Figure 6.3), though which the bulletin board, person directory, group directory, week planner and concept map windows can be opened. By clicking the toolbar legend, the bar can be oriented horizontally or vertically, according to user preference.

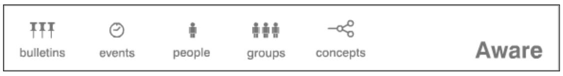


Figure 6.3 – Floating Toolbar

6.2.5 Bulletin Board Window

The Bulletin Board window (Figures 6.4 and 6.5) allows users to compose, post, view and edit message bulletins. It contains three interface panes: a bulletin boards pane where bulletins directed to the user can be viewed, a composition pane from which new bulletins can be posted and a posted bulletins pane where bulletins that the user has ownership of can be revised or deleted. By default, a list of new bulletins is shown; a menu allows access to each of the user's bulletin boards that contains messages. In addition to standard text bulletins, the bulletins board view displays flyers auto-generated by the system to promote an event or group and human-authored media bulletins. An icon through which details of the group, event or media output referenced can be accessed is appended to the displayed message.

The 'New Bulletin' icon offers the user the option of creating either a personal bulletin or a group bulletin on behalf of an `InterestGroup` of which they are a member. The user is guided through the process of authoring and targeting the bulletin by a series of tabbed interface panels. The bulletin message is composed and titled in simple text entry fields. Currently the client allows only plain-text standard messages to be created, but the platform supports any content that can be expressed in ASCII form, so text formatting options could be added in the future. The addressing options offered depend on whether a personal or group bulletin is being authored. The personal bulletin may be targeted in respect to the author's own interests, their group memberships and events that they have attended or plan to attend; a group bulletin may be targeted in respect to interests that group members contribute to the group through their declared involvements, the membership of the group or related groups and events that group members have attended or plan to attend. Each bulletin may be set for immediate display, or timed relative to an event; again, for personal bulletins, this is an event that the user has attended or plans to attend and for a group bulletin an event that a group member has attended or plans to attend. The display period of the bulletin is limited by the client to not exceed 6 months, though the bulletin can be manually reposted to extend its display period or withdrawn before it has expired. When the new bulletin is posted, the client interacts with the server to create and link the bulletin itself and requisite target profile and time interval facts. Immediately this is completed the rules populate the target profile with identifiers of matching interests, involvements and diary entries and append the identifier of the new bulletin to the corresponding bulletin board slots.

The ‘My Bulletins’ pane allows standard bulletins posted by the user, or on behalf of a group of which they are a member, to be revised or retracted. Revisions are carried out in a pre-populated set of the panels used for authoring messages – thus the content of the message can be changed or the target profile to which it is addressed altered. In this way, a simple mechanism is provided for group bulletins to be revised by group members before they are made publicly viewable. A draft group message can be authored by one member of the group and initially made visible only to other group members. Other members can use the kiosk interface to revise the message to make any textual changes that they wish, then when the bulletin is finalised, the target profile is revised to reflect the intended recipients.

Figure 6.4 Bulletin Board Window: New Bulletins board and Personal Bulletin creation tabs

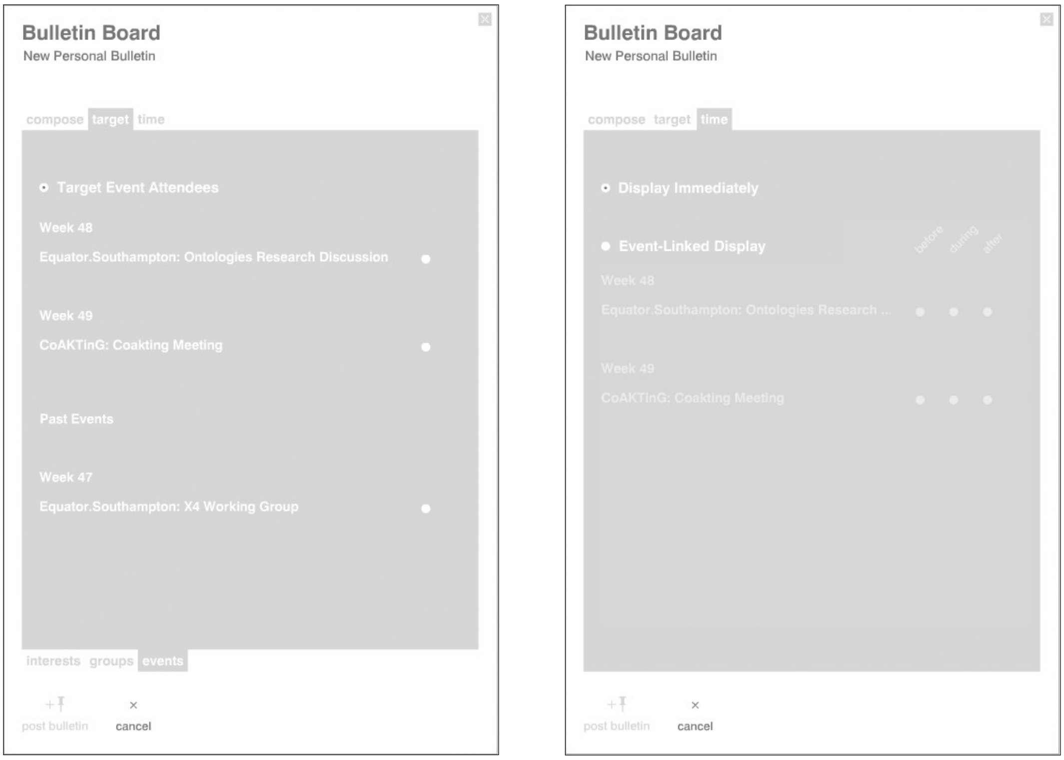


Figure 6.5 – Bulletin Board Window: Personal Bulletin creation tabs

6.2.6 People Window

The People window, show in Figure 6.6, provides a conspectus of the system user base. Naturally the platform allows a straightforward sorted text list of users to be produced. In this client, though, a pictorial representation has been chosen as better encouraging serendipitous discovery. As the user rolls over the array of person icons to locate a particular individual, pop-up annotation identifies their name and role; an individual’s profile is obtained by clicking their icon and no additional effort is required of the user for them to view the profiles of other individuals that catch their attention. The icons are coloured according to the assumed relevance of the individual to the user. Expertise sets are retrieved from the server for each of the user’s declared interests and composite lists of individuals with expertise or interest relevant to the user’s own interests compiled. Icons of those with expertise are coloured pink, those with interest blue. Other individuals that are current collaborators with or declared friends of the user are represented by a strong green icon and the remainder by a pale green one. This pattern of colours, which will be subject only to gradual change over time, will also serve as an orientation device by which the user can learn to readily identify the alphabetic area in the list that they require.

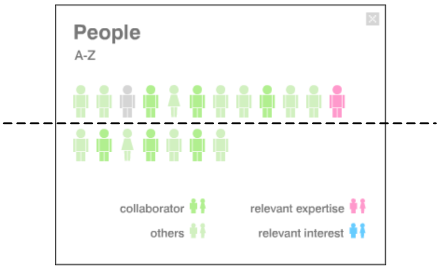


Figure 6.6 – People Window

6.2.7 Personal Profile Window

In contrast to the visual presentation adopted by many of the client windows, the Personal Profile demonstrates a more traditional, CV-style overview of an individual's professional roles, interests and involvements. These are explained in the individual's own words, taken from the `descriptiveText` slots of the corresponding stored facts. Historic formal group involvements are included, sorted according to the year in which they ended. If the individuals has authored any `MediaItems` that the user is entitled to view, an icon by which these can be accessed is included in the corner of the pane.

6.2.8 Group Directory Window

The Group Directory window (Figure 6.7) allows the user to browse existing groups, manage their involvement in informal watercooler groups and create new watercoolers. Three directory lists are available; showing all current groups, historic formal groups and groups of which the user is currently a member – a 'my groups' directory. These lists are selected from a menu that is accessed by clicking the 'groups' icon in the corner of the window. Each group's entry in the list comprises the group title and a descriptive abstract – the first paragraph of text in the group fact's descriptiveText slot.

Those watercoolers included in the current groups list of which the user is not a member have a 'join' icon appended to their description. Through this, access is gained to a 'get involved' tab, with which the user can introduce themselves to the cooler membership. They simply title and describe briefly their planned involvement, then select one or more of their interests that they will bring to the group. When the 'join 'cooler' button is pressed, the system creates the necessary InformalInvolvement and TimeInterval facts. New InterestTargets are also created corresponding to the interests that the user has chosen to contribute; references to these are added to the *advertiseTo* TargetProfile to broaden the range of individuals to whom the group will now be advertised. Finally the group's relationship with other groups will be reassessed to take account of the newly contributed interest(s).

Creation of a new watercooler is only marginally more complicated. The user must first title and describe their new group, then, using the same tab provided to join an existing group, outline their planned involvement in it. When the tab is submitted the system again constructs and links the necessary facts and assesses the new group's relationship with those already extant in the fact base. Once an individual is a member of a watercooler, they can amend their involvement by locating the group in the 'my groups' directory list and selecting its 'edit' icon, returning them to a pre-populated 'get involved' tab. A 'leave' icon is also provided allowing instant withdrawal.

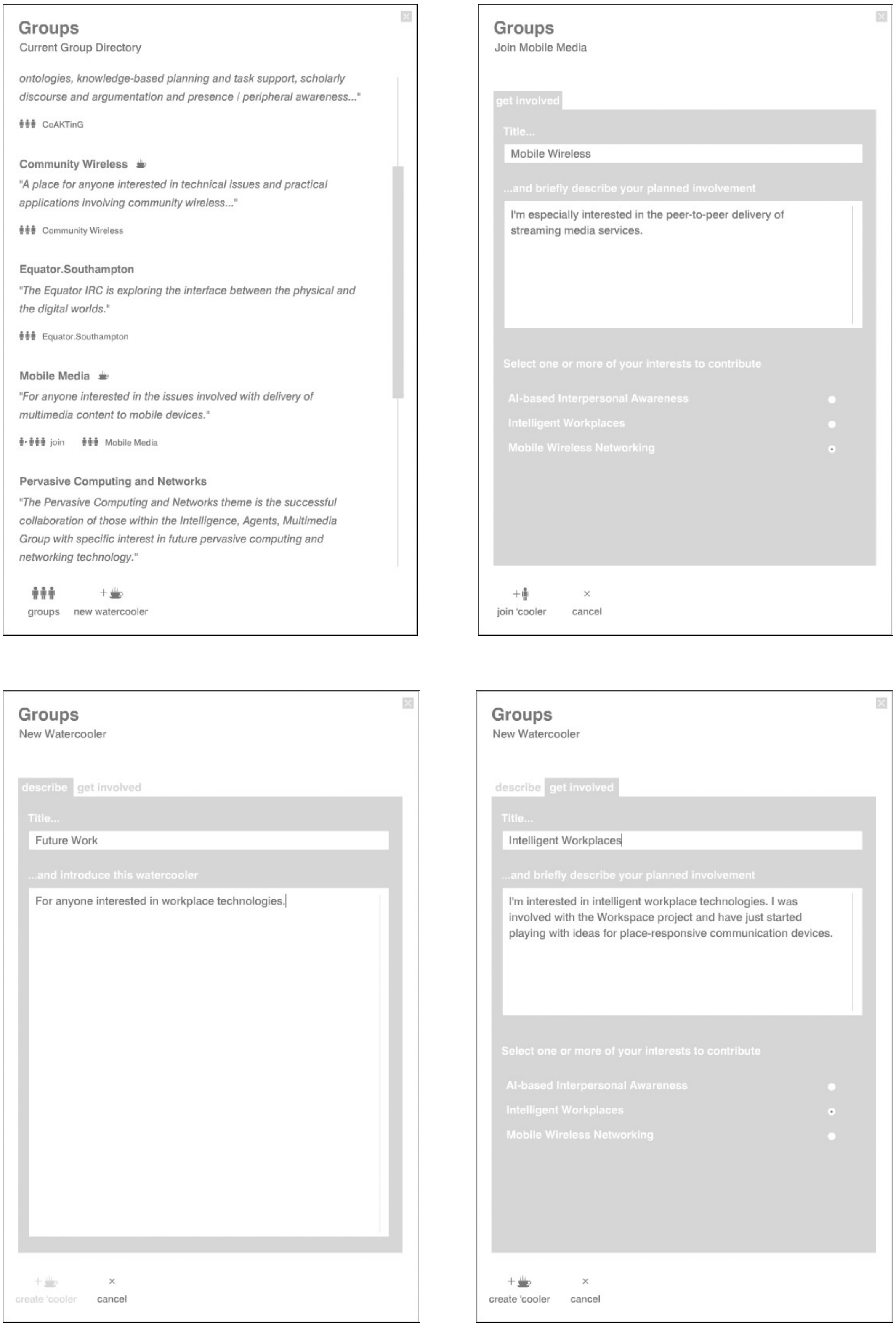


Figure 6.7 – Groups Window

6.2.9 Group Profile Window

The Group Profile window presents a concise summary of a group's current state. A textual description of the group is given, taken from the `InterestGroup` fact's human-authored `descriptiveText` field. Below this, the current membership is presented as a block of person icons. These represent the gender of the individual and the nature of their involvement with the group; administrative support staff icons – i.e. those of individuals that are associated with the group by an `AdministrativeInvolvement` – are coloured a paler shade of green than those of other group members. As with the People window, individuals with interests or expertise relevant to the user's own interests are highlighted, allowing boundary members to be identified immediately. Rolling the cursor over a person icon displays an information box that presents the individual's involvement with the group in their own words, taken from the `descriptiveText` slot of the corresponding `Involvement` fact. If the user is eligible to view any group media items (i.e. `MediaItem`s that reference in their <u>`authors`</u> slot a `ContributorInvolvement` that is associated with this group), an icon through which a list of group media items can be accessed is shown at the foot of the window.

6.2.10 Group Neighbourhood Window

The Group Neighbourhood window (Figure 6.8) provides a convenient summary of groups that are 'in the same space' as one whose profile has been viewed. *`InterestGroupPairScores`* facts that relate a group <u>*`to`*</u> the profiled group are retrieved from the server. Thresholding of relationship strengths is applied to discard those groups whose strongest link is less than 60% of the most strong group relationship returned. Significant relationships are then tabulated to present the user with those groups considered a significant 'supporter', 'neighbour' and/or 'consumer' of the profiled group. The separation of the main Profile window and this function allows a user to maintain a peripheral awareness over time of groups that are relevant to one of which they are a member, without their workspace being unnecessarily cluttered.

Figure 6.8 – Group Neighbourhood Window

6.2.11 Media List and Media Viewer Windows

The Media List window details `MediaItem`s authored by an individual or group. They are arranged under the title of the publication – `the MediaOutput` – that entitles the user to view them (*multiple viewable* `MediaOutput`*s may reference a single* `MediaItem`*, in which case it will appear more than once in the list*). Publications are sorted by date of issue. Though the ontology developed does not currently support explicit chaptering within publications, the ordering of `MediaItem` references in a `MediaOutput`'s <u>`mediaItems`</u> slot determines the order in which they will be shown in the window. Only those publications of which the user satisfies a `TargetProfile` referenced in the `MediaOutput`'s <u>`visibility`</u> slot are returned to the client by the server. Therefore only those documents that the user is permitted to view will be listed; the user is not frustrated by references to documents to which they are not allowed access. Furthermore they are not required to remember multiple passwords to retrieve documents targeted at different groups as their identity, proved by their iButton, is matched directly against the access criteria specified by the publisher of the document. Besides the obvious convenience afforded the user, with this kiosk-based client there is a secondary security benefit in not requiring private passwords to be keyed in a public space.

The Media Viewer window allows the user to read, view or listen to media listed in the Media List window. As the media will be presented within the Flash client environment, it must be available in a format supported directly by Flash, or converted into a standalone movie clip that can be dynamically loaded by the client. Fortunately most common media formats can be accommodated. Popular multimedia formats, including ones for streaming video and audio, are natively supported by the Flash environment. A freely available command-line application, PDF2SWF, is employed to translate Acrobat-format documents to standalone Flash movies, complete with navigation controls styled to match the client user interface. Commercial applications are available that can translate any Windows-printable document to pdf format, allowing other file types such as Word or PowerPoint to be displayed in the viewer.

6.2.12 Community Window

The subject Community Window (Figure 6.9) exploits the platform Expertise Identification function to provide a visualisation of individuals that have an interest related to a given interest or topic. Each individual is depicted by a person icon, coloured according to their estimated degree of relevant interest. A pink figure denotes expertise, a deep blue one interest and a pale blue figure possible interest. The expertise sets returned by the platform include the interests and group involvements that informed its decision to include each individual in the model. This information is presented to the user in the form of a mail-merge generated natural language description, accessed by rolling over the individual’s icon. Though this computer generated text is differentiated from that authored by humans by being presented without italicisation and quotation marks, the distinction is intentionally subtle and it can be seen that the ontology used by the platform allows the two forms to be used interchangeably in a client; if, for example, the user provides a lengthy description of their involvement with a group and a client application for a mobile device calls for a much more succinct piece of text, this can be generated by reference to the ontology and presented in its place. Where boundary members are identified – individuals that have an interest relevant to the community *and* share an area of interest with the user – these are grouped separately in the same window, under the title of the user's corresponding interest. Here, the rollover information box also describes the individual's relevance to the user's interest.

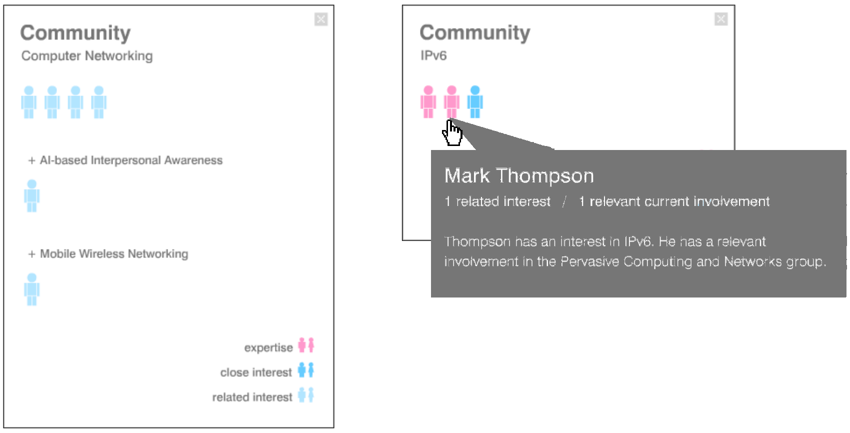


Figure 6.9 – Community Windows

6.2.13 Connections Window

The Connections Window provides a single-click tool that suggests the most promising opportunities for interaction with a colleague, a stranger within the organisation or a group. It is accessed via a 'connections' button, displayed on each Personal Profile and Group Profile window. Two forms of connection are considered: chains of acquaintance identified by the system's social network function and opportunities for direct contact presented by social gatherings. Direct contact opportunities are determined by first identifying forthcoming `SocialGatherings` that the target individual, or one of the target group's members, has indicated through a `DiaryEntry` that they will be attending physically[7]. Those gatherings that are not open to the user are disregarded, as are those that temporally overlap another event that the user is planning to attend. The

[7] As virtual gatherings do not typically present opportunities for informal private discussion, only planned attendance at `PhysicalVenues` is considered.

contact opportunities that remain are then assessed to identify the best candidates. Each opportunity is weighted according to the immediacy of the event – naturally the sooner the event, the more highly it is scored. Events that the user already intends to join are favoured; in addition, if the target is a group, social gatherings hosted *by* that group are weighted more strongly than others that a group member happens to be attending, as they clearly offer greater opportunity to 'connect' with the group as a whole.

From the social chains and any direct meeting opportunities identified the two highest scored connections are selected and displayed in the window, both in iconic form and as natural language explanations, generated using simple mail-merge techniques. In the event that no connection can be found, a suitable text message is shown. Example Connections Windows for different target types are shown in Figure 6.10.

Figure 6.10 – Connections Windows

Figure 6.11 – Events Window

6.2.14 Events Window

The Events Window shown in Figure 6.11 offers a week-by-week visualisation of future social gatherings open to the user. A familiar timeline format is used, with each event represented by a titled bar. This allows the user to identify at a glance relevant events that they are free to attend. A colour-coding scheme is used to denote events that are in the user's diary, those that they have confirmed they will be joining virtually and those that they intend to physically attend. Events that the user has not placed in their diary but that have a venue advertised to a `TargetProfile` that matches them (*and thus are deemed most likely to be of interest to the user*) are also highlighted. The description and location of each event are available in a rollover captions attached to the event bar. By clicking on an event bar, the user can add the event to their diary for future reference, announce their intention to attend, virtually or physically, depending on the venues open to them, or remove the event from their diary, cancelling any planned attendance. With each click, a corresponding `DiaryEntry` fact in the system is created, modified or retracted. Currently users are not prevented from stating their intention to attend multiple events simultaneously, or alerted if one event that they plan to attend is retimed such that it clashes with another. It is envisaged, however, that this would be addressed by the scheduling mechanism proposed in Chapter 8 as a future development.

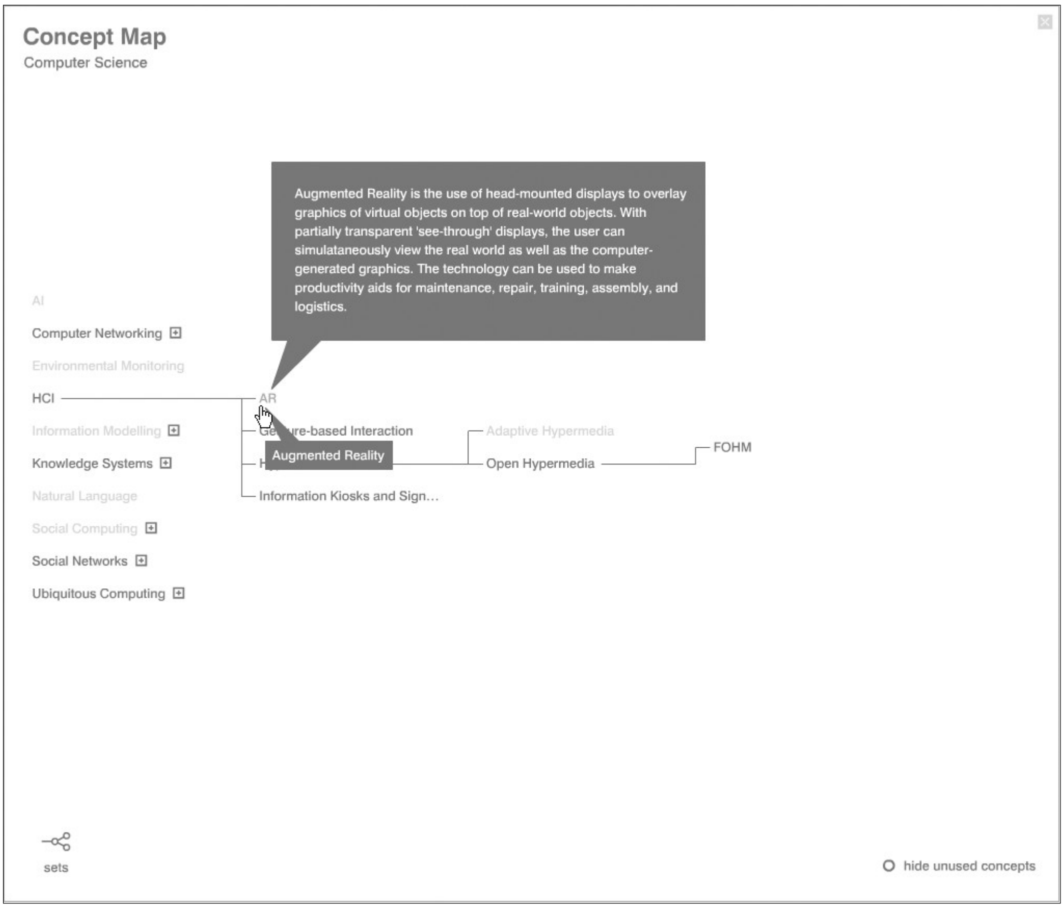


Figure 6.12 – Concept Map Window

6.2.15 Concept Map Window

The Concept Map provides a glossary, allowing the user to browse each of the topic taxonomies held in the system through an easy-to-navigate tree structure (Figure 6.12). Rollover captions display any descriptive text associated with a topic. By visualising those topics that have not attracted interest within the organisation, it can also suggest areas for future work; as argued by one of the Volvo employees that Lindgren et al. surveyed, it will "*be able to show shortcomings, missing competences… you could find areas that were neglected*" [Lindgren03]. Topics for which the system has created *TopicInterestPairScores* facts with at least one relationship scored over 25% are highlighted and can be selected to display the corresponding Community Window. A checkbox option allows those topics not associated with any declared interests to be hidden. This will be of most use where only a small portion of a taxonomy is relevant to the organisation.

6.3 Summary

A set of platform usage situations within a research organisation environment has been identified and corresponding client application specifications developed. One of these client applications, a public kiosk interface, has been implemented to demonstrate how functionality provided by the Aware platform can be packaged as a coherent and intuitive application. In the following chapter, the suitability of the platform for deployment in a very different environment of innovation will be considered as a component of the evaluation.

Chapter 7

Evaluation

In this chapter, the Aware platform will first be judged against the requirements for an innovation support platform identified in Chapter 4. Having demonstrated in the previous chapter that the platform can support an intuitive kiosk-based application for use within a single innovation-led organisation, the flexibility of the platform is here tested by consideration of a potential deployment in a very different environment of innovation, the organic farming community. Finally, the platform will be contrasted with a conventional query-led architecture and the pros and cons of maintaining an active organisational model will be examined.

7.1 Satisfaction of Requirements

In Chapter 4, three key challenges that an innovation support platform must address were identified: awareness, community and communication. More general requirements, that the platform provide timely information and support a range of client device types, were also identified. In this section, the degree of support offered by Aware to the challenge areas will be considered; the other issues will be examined later when the platform is contrasted with an alternative software architecture. Some attempt has been made to segment the following discussion by challenge, though, as these were shown to be interdependent -- and have been treated as such in the development of the platform -- there will inevitably be a degree of overlap.

7.1.1 Awareness

Aware supports awareness of individual and group activities and of the relationships between individuals and groups. It is scarcely believable that one individual within an organisation would introduce another as someone that they had co-authored a document with but neglect to mention that they are currently collaborating on a number of informal internal projects, yet that situation is echoed in the social relationships identified by social and professional network evaluation tools such as those surveyed in Chapter 3 that base their relationship models on ontologies intended for information exchange between organisations or information 'harvested' from public web sites. Because Aware is specifically designed to support awareness from *within*, for individuals inside the organisation, or community, that a particular deployment has been set up to serve, it supports social network visualisations that detail connections that are both current and immediately exploitable, rather than relying on co-authorship of publications that might have been written several months ago and whose authors might not actually have ever physically met each other. Simple, concise textual explanations of the social links that users can use as tokens of social introduction are easily generated because the ontology that Aware uses is structured to model the natural relationships within organisations rather than obscure them.

Topic experts, similarly, are straightforwardly identifiable. In contrast to the tools surveyed, in which the user must 'guess' an individual most likely to be associated with the subject in question, Aware allows those sharing interest either in individual topics or with users' declared interests to be identified *and* uniquely highlights those all important individuals that have 'common ground' with the user, sharing one or more of their own interests. This feature is also exploited to provide visualisation of boundary members within Communities of Interest.

Awareness of Communities of Interest and events is also well supported. The automated creation of targeted flyers to advertise water coolers and events to relevant individuals provides a natural and unobtrusive means to place them within the user’s everyday ‘line of sight’. The more subtle method chosen to enhance awareness of formal groups, within a dedicated ‘Group Neighbourhood’ tool, is appropriate as these groups offer a less immediate opportunity for interaction. Though the platform does not currently share the sensitivity to physical proximity of Inter-Personal Awareness Devices, its active conceptualisation and the inclusion in the ontology used of a model of physical location that can be linked to an individual ensure that such functionality could be added at a later date with minimal effort.

7.1.2 Community

The most significant practical concern identified regarding existing solutions to support Communities of Interest was that they inadvertently deter speculative creation of new communities. The model of community and community involvement that Aware adopts minimises both the psychological commitment and actual effort required of an individual wishing to launch an embryonic community – that is, a water cooler of which they are initially the sole member. Having titled and described their water cooler, they need only identify one or more of their declared interests to contribute. Because the relationships between communities are inferred by the system, not dictated by a user-managed hierarchy, they are not required to modify some public community map or explicitly place their community relative to others that already exist. The water cooler will then automatically be promoted to individuals to whom the *system* judges it will be of interest (*though its creator is, of course, free to advertise it more widely with a manually authored bulletin*). Should a new water cooler fail to attract interest, it is easily later withdrawn by its creator without their ‘loosing face’; simply by retracting their involvement they cause the then-empty community and its promotional flyer to be removed from the system.

The involvement model implemented also ensures that Aware respects the fine distinction that has been made between community *lurker* and *boundary* *member*. It scuppers the would-be lurker who wishes to leach from the community but remain invisible and contribute nothing – they must declare an interest in *something* in order to join – while ensuring that boundary members that have a valid peripheral interest in the community and who are an essential element of an innovative environment are readily identifiable.

Communities of Practice are recognised as presenting a particular challenge in an innovation-led environment: they form essential human repositories of tacit subject knowledge, but can be inward looking and impenetrable to non-members. Aware provides a pragmatic solution that exploits its unified conceptualisation. If a Community of Practice's members wish it to be explicitly represented, with an identity and defined membership, it can be treated in the same way as a Community of Interest, modelled as either an open water cooler or a formal group. This affords them collective public visibility, system support for social gatherings and the ability to direct bulletins to the community membership. Regardless of whether the community is explicitly represented in the system, though, the knowledge that is gained collectively in it is accessible to outsiders. As individuals, its members contribute their expertise to other communities that they are involved with (that is, they *practice*), so are included in the Expertise Identification provided by the platform, along with any non-members that also have relevant interest or expertise. The effectiveness of the Community of Practice is not compromised; to its membership, value is gained from their collective identity, but to outsiders, this is not important – for them, the community is simply a pool of knowledgeable individuals. Countering their potential insularity, Aware's support for targeted bulletins and active promotion of events ensures that Communities of Practice are regularly exposed to external ideas, both from non-members with similar expertise or interest *and* those carrying out work that supports or exploits their field.

7.1.3 Communication

Of the three areas to be supported by the platform, communication stands alone in being partially satisfied by existing mature technologies. Accordingly, no attempt has been made to duplicate the functionality provided by Instant Messaging and similar real-time person-to-person communication technologies, though the platform is quite capable of interfacing *with* them. Instead, attention has been focussed on relieving overuse of email by providing a more appropriate medium for broadcast messages and on encouraging face-to-face communication through social gatherings.

To avoid the issues that plague email simply being transferred to a new medium, two problems had to be addressed: the receipt or continued display of messages that are of no relevance to the user and the non-receipt of messages that have gained relevance to the user since they were posted. In supporting a target-based addressing mechanism, the platform provides users with a straightforward way to convey their message to relevant individuals, regardless of whether they happen to know who those individuals are. As Goecks identified [Goecks02], messages may still become irrelevant simply because the events that they refer to have passed. A fixed window of display can be specified for a bulletin, allowing the length of time for which it will be shown to be limited. Goecks' concern is better addressed, though, by the platform's ability to support timing of bulletins relative to a specified event. This temporal relationship is evaluated dynamically, ensuring that if, after the bulletin has been posted, the event is then rescheduled, it will still be displayed when its author intended.

Because the 'delivery' of bulletins, too, is performed dynamically, Geyer's [Geyer02] suggestion that messages be available to "*late joiners*" is also satisfied. If, for example, a number of bulletins are targeted at attendees of a forthcoming event, they will immediately become visible to individuals that have already indicated their intention to attend the event. However, if another individual later indicates that they too plan to attend the event, the bulletins then become visible to them, even though they had been posted earlier. As well as automatically promoting events through flyers and offering a convenient mechanism for users to post bulletins related to them, Aware further encourages face-to-face communication by its unique consideration of future social gatherings as a potential link in a social network, a feature that illustrates, perhaps better than any other of the platform's services, the value of an integrated treatment of awareness, community and communication.

The popularity of URL recommendations that was highlighted by Goecks has been recognised with the provision of a dedicated bulletin type and, following Fischer and Ostwald's [Fischer03] observation of the importance of documents in grounding shared understanding, this concept has been expanded to allow published documents and other media objects represented in the conceptualisation to be similarly promoted. A significant improvement on the functionality of NuggetMine and other such tools is the inherent ability of the structure of these specialist bulletin types to minimise the potential 'nuisance factor' of such postings. In addition to the targeting mechanism common to all bulletins, the use of dedicated slots in which the address or identifier of the recommended item is referenced ensures that users need not be swamped with multiple recommendations to the same item; all such recommendations can easily be consolidated by a client application as one message or group of messages. Conversely, users need not hold back from posting potentially useful recommendations for fear of causing annoyance.

7.2 A Flexible Solution

The kiosk application that has been developed vividly illustrates the ability of the Aware platform to support sophisticated, multifunctional clients for deployment within what might be considered a traditional environment of innovation. In this section, the capability of the platform to support alternative client types and usage scenarios is demonstrated; a lightweight mobile client for meeting support is described and the potential of the platform to support innovation beyond the corporate and academic worlds is explored.

7.2.1 ContactAware

The framework developed in Chapter 4 identified the need for the platform to support simple single-function clients on mobile devices. In meetings, it is often desirable to obtain specialist advice from colleagues not present. ContactAware is a mobile phone-based Aware client application that allows its user to quickly identify colleagues with expertise in a topic that has arisen and contact one so that they can remotely participate in the meeting.

Developed using J2ME embedded Java, ContactAware provides access to Aware's expertise identification function (described in Section 5.3.1). The topic taxonomies installed in the platform are parsed to construct a predictive-text-enabled input box, in which the user specifies the subject in which they wish to find an expert. The expertise identification function is called with the identifier of the chosen topic and the returned set matching individuals presented as a list of names, sorted in descending order of relevance. Using the phone's cursor pad, a name in the list can be highlighted. The phone's soft buttons then allow the user to place a call to the highlighted individual at one of the numbers linked to them in the platform's knowledge base, or obtain a textual description of their relationship to the topic of interest; this is constructed using the same mail-merge routine that was employed in the kiosk-based client (see Section 6.2.12).

ContactAware caches the platform's taxonomies locally, only retrieving from the server new or newly modified topic structures. The remaining exchanges between client and server typically involve only a few hundred bytes per session, so is viable to use charged-per-byte GRPS connectivity, allowing the application to be used where there no access to local network infrastructure. ContactAware can therefore serve its user beyond the walls of their organisation, helping in offsite meetings to overcome the physical isolation between them and their colleagues. ContactAware also illustrates how the platform's functionality can usefully be integrated with that already provided by a client device – in this case, the device's existing facility to make telephone calls is augmented by the platform's ability to identify individuals that might have information of value to the user.

7.2.2 An Alternative Application Domain

Innovation is not the sole preserve of research laboratories and design studios. As the platform is not anchored to any particular topic taxonomy, it is domain-agnostic and should be applicable to a wide range of innovation-led environments. The amount of organically managed farmland in the UK has increased twelve-fold over the past six years, driven by what is now a billion-Pound UK market for organically produced food[1]. Farming successfully without synthetic chemicals or genetically modified organisms demands constant innovation, both in agronomy and animal husbandry and in marketing and business models. Though innovation in organic farming (OF) is supported by a number of formal government and charitable research establishments, "*contrary to suggestion that most of the research on organic agriculture is done at public institutions… much of the innovative structure of organic agriculture may be developed on farms or in a learning-by-doing setting*" [Faber03].

OF is a highly attractive domain for the Aware platform; the communities and personal networks that Aware exposes are essential to the success of farm-level research. The Organic Farming Research Foundation[2], which encourages organic farmers to form local 'research clubs', reports that 83% of US-based

[1] www.soilassociation.org

[2] www.ofrf.org

organic farmers cite informal farmer-to-farmer networks as an information source. Academic support is provided by Milestad and Darnhofer, who state that "*the resilience of the organic movement depends on its connection, communication and inter-actions…* [to] *allow knowledge systems to accumulate and be transferred over time.*" [Milestad02]. The potential value of a software tool that can assist this connection, communication and interaction (*which surely equates with awareness, communication and community*) is particularly apparent when one considers the consistently high proportion of newcomers in the OF community – currently around 7% of organic land is in the first year of conversion – that has resulted from the rapid growth and large potential for future growth of this domain.

Though the ontology that Aware implements has been documented in reference to an institutionalised environment of innovation, it accommodates well this alternative scenario. The interest group conceptualisation adopted is flexible enough in its support of location, ownership, involvement and funding that it can model – and identify communities of practice and expertise from – the gamut of individual and collaborative activities associated with OF, from office-based marketing initiatives to an agricultural trial, located on a farm and carried out by its personnel, but funded by an outside agency. The RDF-lease model presented will allow of information from such agencies to be integrated easily into the domain conceptualisation. Besides supporting the promotion of interest groups, the provision in the ontology for bulletins targeted according to interest, involvement and location will find many other applications in this domain. It offers a mechanism to solicit and share informal tips (a *typical example of which, from the Goodbugs*[3] *mailing list, is reproduced in Figure 7.1*) without requiring knowledge of and subscription to multiple mailing lists. It also provides an opportunity to disseminate news of trial results, the media output publication structure allowing results of ongoing work to be presented as a research diary.

[3] http://web.agrsci.dk/plb/iobc/goodbugs-l.htm

```
I am looking after information about Contarinia
nasturtii, the swede midge.

Does anybody know a solution to control this pest on
broccoli or cauliflower with biological methods ?
```

Figure 7.1 – A typical request for advice posted to an organic pest control email list

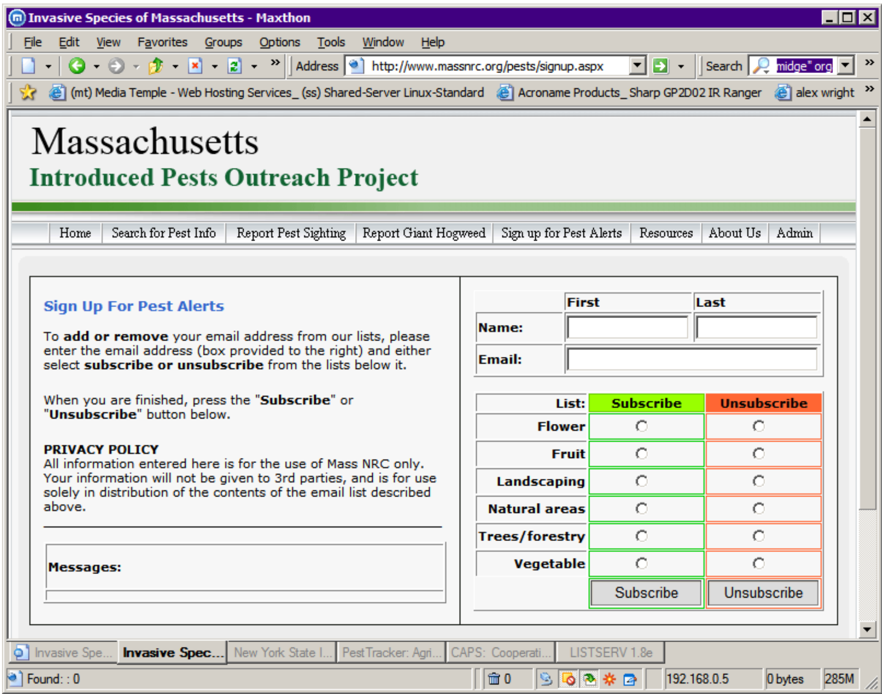


Figure 7.2 – Pest Alert list signup page

As Aware bulletins are modifiable after posting and can be retracted at will, they present an ideal medium for the delivery of pest alerts and meteorological warnings; currently such messages are shared via broadly-targeted websites and email lists (Figure 7.2)[4], with the inevitable consequence that a subscriber receives many irrelevant notices. Timed targeted bulletins offer in this scenario an attractive means of publicising relevant broadcast media items. The BBC's daily Farming Today[5] programmes, for example, are archived online and

[4] www.massnrc.org/pests/signup.aspx

[5] http://www.bbc.co.uk/radio4/news/farmingtoday/index.shtml

the URL bulletin provides an opportunity for programme producers or audience members to selectively promote programme items[6]. The support offered for professional gatherings might not appear so useful, yet OF is in reality characterised by frequent open knowledge-sharing events held at farms and supported by a busy programme of formal seminars and conferences.

Given the primacy of topic relationships in Aware, a prerequisite of deployment of the platform in this domain is the availability of a widely agreed taxonomy of agricultural concepts. Fortunately, such a taxonomy exists; AGROVOC[7] is a hierarchical, multilingual agriculture thesaurus, maintained by United Nations Food and Agriculture Organization. Freely available in machine-readable RDFS format, it contains over 16,000 terms relating to agriculture, fisheries, forestry and the environment and is the standard reference for OF terminology [Slijkhuis96]. The ability of the platform to support multiple taxonomies allows a large source taxonomy such as this to be presented to the user as a series of smaller, easier-to-navigate sub-taxonomies.

While the platform needs no alteration to support this domain, it is not suggested that the client suite proposed for an institutional environment would be used. The situational types involved will differ somewhat from those identified in the previous chapter; there is potential for deployment of public access kiosks in agricultural centres, but it is envisaged that the majority of interaction with the system would be through home-office desktop PC clients that combine the functionality of the four client types previously described. There will also be cosmetic differences in the presentation of information returned by the platform. The water cooler metaphor, for example, has little resonance in this scenario, so

---

[6] In the final chapter, the use of the platform to selectively promote video or audio segments is further explored as a means to distribute meeting and seminar material within organisations.

[7] http://www.fao.org/agrovoc/

while the concept of an open knowledge exchange group remains valid, a more appropriate representation would be used in the client user interface.

A final consideration is the technical requirements that the system has of its users. The architecture of thin-clients coupled to the server via a lightweight XML protocol was developed primarily to allow the use of resource-constrained pervasive computing devices within organisations, but also serves to ensure a satisfactory user experience in the more modest technical environment that this scenario offers. A standard office-grade PC and low-bandwidth dialup connection which will offer performance approaching that that would be achieved from a top-flight machine on a local network. It can be seen, then, that although the design of the platform and much of the research that informed it were driven by the needs of individual technology-led organisations, the approach adopted is flexible and generic enough that it can equally well support a distributed and ostensibly non-technical environment of innovation.

7.3 Comparison with a Query-led Architecture

An architecture has been presented in which a palette of communication, community and awareness services is supported by an active organisational conceptualisation, of which an important feature is a dynamically maintained network of topic, interest and interest group pair evaluations. A useful comparison can be made between this approach, implemented using the JESS inference engine and that adopted by ‘static’ database- or triple store-based systems, in which evaluation of the organisational model is instead performed on an as-required basis, by individual client services.

In Chapter 4, we briefly introduced CS AKTive Space [Schraefel04], a triple store-based directory of UK computer science research. Like Aware, CS AKTive Space contains a central domain conceptualisation, or ontology, to which a number of client services for the exploration of this conceptualisation are linked; it, too, allows the activities of organisations and individuals to be explored. One of the services supported is the ONTOCOPI network visualisation, examined in Chapter 3. Superficially, notwithstanding structural differences in the domain conceptualisations used, ONTOCOPI functions in a very similar way to the Community of Practice visualisation in Aware. It develops a set of numerical evaluations of interpersonal relationships between individuals represented in the domain conceptualisation, which are used to compile a set of individuals judged most closely involved with a given person.

The important operational difference between the two tools is in the way that the relationship evaluations are achieved. While the numerical evaluations that the Aware tool uses are a component of the platform's central conceptualisation, in CS AKTive Space / ONTOCOPI, no common evaluations are maintained. Instead, when a user invokes the ONOTCOPI component, it retrieves relationships from the domain conceptualisation and then evaluates them as required. In Aware, then, evaluations of significant conceptual relationships are actively maintained within the domain conceptualisation to mirror any changes made to the conceptualisation and available to all external client services, whereas in an architecture of the type that CS AKTive Space represents, the evaluations are performed *by* the client services at their behest.

Aware's Community of Practice tool returns a requested community almost instantaneously, whereas ONTOCOPI users experience a short delay while the interpersonal relationships retrieved from the triple store are evaluated. In this instance, though, it could be argued that the elimination of this delay is not significant; even allowing for the more complex network of evaluated relationships that the Aware tool requires to produce its Community of Practice model, the value of the service will override any minor inconvenience incurred, particularly as the user will most likely request only one community to be identified in a session. The Aware platform, though, exploits evaluated relationships to support a range of different services. Consider the processes that are involved in the delivery of a bulletin that has been targeted to consumers of a particular interest group's work. A set of topic type relationships is formed, based on the relative positions of topics in their taxonomy. Topic support relationships are evaluated, based on co-occurrences of topic references in interest profiles. Using these sets of topic relationships, pairs of interests are evaluated according to ten different relationship criteria, to identify three types of interest relationship. From these, interest group pairs are evaluated, finally allowing sets of interest group neighbours to be identified, against which the bulletin's address target profile can be matched. If the bulletin is timed to be shown relative to an event, its display time will be resolved by reference to the time period of the event itself. Only now can it be determined to which bulletin boards, if any, the bulletin should be posted.

This bulletin might be one of a thousand or more current group-addressed bulletins in the system; to provide the intuitive bulletin addressing system offered by Aware using a query-based architecture, the evaluation described would have to be performed for each one of these bulletins, every time a user views their bulletin boards. A similar process would have to be carried out to determine whether the user is eligible to attend a particular event when displaying event listings. With a large topic taxonomy and organisational conceptualisation, the amount of computation required would render client applications virtually unusable. Compared with the query-based approach exemplified by CS AKTive Space, then, the advantages of Aware's dynamic evaluation model can be summarised as follows:

- Temporal distribution of evaluation, allowing the system to remain responsive while maintaining complex evaluated networks. Individual relationships are re-evaluated whenever there are changes made to the conceptual model that could alter their strength, therefore there should be no circumstance where a very large number of relationships has to be evaluated at once.

- Reuse of evaluation between services, ensuring that resources are not wasted needlessly re-evaluating unchanged relationships.
- Transference of the burden of computation from client services, which may be located on resource-constrained devices, to the server machine where adequate resources can be ensured more easily. This also ensures that as the conceptual model grows, the platform resources can be scaled to accommodate it without requiring the multitude of client devices that access it to be upgraded.
- The platform is able to respond *proactively* to changes in the conceptualisation. It can react both to direct relationships indicated by the conceptualisation and new or modified evaluated relationships. This allows the platform to dynamically refresh content showing on client interfaces, such as bulletin board listings, obviating the need for the client to repeatedly poll the server. It will also facilitate proposed future developments of the platform outlined in the next chapter, such as push-delivery of notifications to simple read-only client devices.

Because of these characteristics, relationship evaluation can be exploited extensively by Aware services. CS AKTive Space provides services dominated by link traversal and straightforward comparison of concept attributes, with relationship evaluation featuring only in a peripheral service. Aware, by contrast, places evaluated relationships at the centre of its core services; link traversal becomes primarily a means to annotate relationships and to navigate between services.

Ultimately, the viability of a dynamically maintained evaluation network, is dependent on a stable conceptual model characterised by small incremental changes that are infrequent relative to the instances that the evaluations will be used. Were such conditions not to prevail, the performance of the server would be severely degraded by repeated re-evaluation of the network and the benefit of a commonly accessible evaluation model would be lost. Broadly, there are three triggers that cause re-evaluation of inter-concept relationships:

- modification of a loaded taxonomy or introduction of a new taxonomy to the system.
- creation, modification or removal of a user's declared interest profile.
- creation, modification or expiry of a user's declared involvement with an interest group.

The most disruptive change that can take place is modification of the structure of a taxonomy, as addition or removal of a single topic can cause all relationships within its branch to be re-evaluated, encompassing topics that have no practical connection. This will cause the re-evaluation of relationships involving any interests that reference those topics and, in turn, of relationships linked to interest groups to which any of these interests are contributed. Fortunately, though, it can be presumed that changes to taxonomies will be made only infrequently; it is anticipated that taxonomies will be provided by commercial or public organisations, which will consolidate any changes into weekly or monthly update packages. Modification of users' declared interests and group involvements, though causing much more limited disruption to the evaluated relationships, will occur more frequently, yet, even in environments that suffer heavy staff churn, such as universities, it is clear that alterations to interest profiles and group memberships will be massively outweighed by viewings of bulletins, explorations of Communities of Practice and glances at the forthcoming events diary. A user's typical everyday interaction with the platform does not alter the evaluated relationships, it exploits them, time and time again.

While a query-based architecture can replicate some of the functionality that the Aware platform provides, the ability of Aware to fulfil *all* three of the key requirements of innovation – awareness, community and communication – is heavily dependent on its use of an active, self-evaluating, domain conceptualisation. It is practicable to maintain such a conceptualisation because the relationships within the conceptualisation have a relatively high degree of stability. Future developments of the platform outlined in the final chapter will also depend on its active core to react immediately to changes in the conceptualisation.

7.4 Summary

The Aware platform addresses the primary environmental requirements of innovation, by supporting awareness, community and communication between individuals with similar or relevant interests. Though the main client scenario developed is intended for deployment on PC hardware within a single innovation-led organisation, it has been shown that Aware is a flexible solution and its services can be accessed by lightweight mobile client devices and are equally relevant to a dispersed, ostensibly non-technical environment of innovation. The active, self-evaluating conceptualisation on which Aware is based has been contrasted with a more traditional query-led architecture, demonstrating that only this active architecture can viably support the full range of services that Aware provides.

Chapter 8

Future Work and Conclusion

It has been shown that Aware provides flexible support for the key environmental requirements of innovation. There are, nonetheless, areas in which it will benefit from further work, to provide additional functionality for the user or ease its integration with other business services. In this final chapter, a number of potential enhancements to the platform and new client applications will be suggested. There are also opportunities to further develop the background research on which this work is predicated; these, too, will be briefly considered. The thesis concludes with an overall summary of the work that has been presented.

8.1 Future Platform Enhancements

In development of the Aware platform thus far, emphasis has been placed on aspects of its functionality that demonstrate the value of its active unified domain conceptualisation. Though it is fully operational and ready for real-world deployment in its current form, there are several areas in which further development will enhance the platform's usefulness or usability.

8.1.1 Automated Scheduling

Aware can dynamically advertise an event to individuals with certain interests or group memberships, but it is the responsibility of the event's organiser to schedule it for a time when such individuals are able to attend. Similarly, it is the responsibility of an individual intending to attend the event to ensure that they do not have any other commitments at that time; the platform will not prevent an individual committing to being ‘in two places at once’. By contrast, automated schedulers, a typical example of which is BBN’s system to time-manage aircrews[1], dynamically programme events according to a set of business or process rules and warn their users if there is an irresolvable problem.

If automated scheduling capability were added to Aware, it could prevent individuals creating clashing appointments – and alert them if a clash were caused later by an event that they plan to attend being rescheduled. More interestingly, having access both to users’ plans and to the dynamically populated event target profiles, it could place events according to when the largest potential relevant audience will be available. When creating a new event in the system, a user would have the option of using a ‘find me a time’ function, that would offer a selection of times when the largest possible proportion of individuals matching their event’s target profile were able to attend.

With automated scheduling, Aware would also be able to suggest alterations advisable to longstanding or repeat events. If a regular discussion group coincided with a guest presentation by a well-known speaker, it is likely that many of those that might consider joining the discussion group would instead attend the presentation. Intelligent scheduling would recognise if a number of those most interested in a subject of the discussion group were unavailable, or if another event with an intersecting target profile had been scheduled for the same time and suggest to the discussion group organiser an alternative time when a larger pool of interested people was likely to be available.

Aware currently uses a relatively primitive temporal model, so some changes to the ontology would be required to implement an effective scheduler. Users would probably expect, for example, that it be possible to constrain the scheduling of an event relative to another event, or within a certain timeframe.

[1] http://vishnu.bbn.com/acs/acs.html

They might also wish to specify an ‘essential core’ of individuals that must be able to attend. Though probably somewhat over-engineered for this application, a potential starting point would be the 13-strong set of temporal algebra relations developed by James Allen [Allen83].

8.1.2 Inter-platform Operation

The platform was first presented as a solution for use within a single organisation. The organic farming scenario presented in the previous chapter then showed how many small organisations could exploit a single Aware server. There are situations, however, in which it will be desirable to achieve some connectivity between client applications paired with *different* Aware servers. Consider a conference that attracts attendees from numerous commercial and academic organisations. Each attendee might possess an Aware client device – a laptop, perhaps, or mobile phone – that is connected to their organisation’s local Aware server. This in itself would be useful, allowing an attendee to identify their organisation’s expertise and activities concerning a topic that catches their attention during a presentation, but it does not directly encourage collaboration between attendees during the event.

Rather than relying on their audience to themselves identify and research salient concepts, though, a talk’s authors might instead wish to embed in it interest profiles to be broadcast at relevant points during the presentation. Attendees might wish to be alerted to others co-present with them that share a particular interest, or actively promote their interest in a subject to attract potential collaborators. Aware’s facilities for organising social gatherings would be of value in promoting and managing attendance of discussion sessions and workshops, its support for targeted bulletins enabling individuals to table media outputs from their own organisation for consideration. A user would also be able to exploit the platform’s social networking function to identify any convenient ‘opportunities’ presented by those sessions and workshops to meet a specified individual from another organisation.

It is likely that a component of this scenario will be a dedicated Aware server used for the duration of the event. A significant challenge will be to facilitate evaluation of interests shared within the event server using the concept support relationships established in a user's own platform; to judge the relevance of another organisation's work in the context of that of their own. A mechanism has been described through which an object resident in Aware's conceptualisation can be tied to a definitive external data source, but no provision has yet been made for objects managed by one Aware server to be temporarily placed and evaluated within the conceptual model of another to serve a single user.

8.1.3 Physical Awareness

The inclusion in Aware's ontology of an (admittedly very primitive) location model invites future inclusion in the platform of some form of physical awareness. Obvious applications of this would be physically constrained bulletins that would only be displayed to a user when they were near or within a particular space (a conference room, for example) and in the creation of an enhanced expertise finder function able to differentiate between physically proximate sources of knowledge and those that could only be accessed through remote communication.

8.1.4 A Web Service Wrapper

The XML interface through which clients are currently able to connect with the platform was constructed on an ad hoc basis, evolving gradually as the demonstrator client itself was developed. It is lightweight, functional and perfectly adequate as a means of enabling the functionality provided by the platform to be showcased. It does, however, have one significant limitation: it requires that any client application that uses it be specifically designed to interact with Aware. Previously it was noted that in a practical deployment Aware clients would be expected to group Aware functionality with other organisational services.

Providing Aware with a web service interface, whereby individual platform services would be both discoverable and self-describing, would development of client applications and insulate them from future modifications to the platform itself.

8.2 Client Applications

In addition to the platform developments suggested, there are some exciting opportunities for new client applications that would require, at most, minimal alteration to the current platform implementation. Where possible, a client application should exploit existing, readily available technology. Undoubtedly the most ubiquitous information appliance today is the mobile telephone. It has eclipsed the PDA, or Personal Digital Assistant, now packaging the personal organiser tools for which these devices were once popular with traditional mobile telephony and messaging. The simplest means by which any mobile phone user can interact with a remote computer system is via SMS text messaging. The platform could easily be linked to an online SMS gateway to provide rapid notification of event cancellations or schedule changes, or allow summarised contact details or social connections to be retrieved. An SMS message may take several minutes to reach its destination though and, regardless, the quantity and quality of information that can be exchanged in 160-character text phrases is limited. Smart phones, which offer window-based GUIs, TCP/IP internet connectivity and ad hoc Bluetooth-based networking with nearby devices, can provide a fuller, more responsive interaction with the platform.

Proximity-based matchmaking and 'buddy alerts' are already a popular consumer smart phone application[2]. A phone-based Aware client could alert a user as they pass a new employee with interests similar to their own, or the venue

[2] Proximity is determined by connectivity being established between the user's phone and another Bluetooth device; this might be another individual's phone, or a fixed Bluetooth access point attached to a building or some other object.

of an event that they might wish to attend. Soon, phone-based applications will have access to fine-grained *absolute* positional information. Motivated by forthcoming US legislation requiring that the origin be identifiable of 911 emergency calls placed from mobile phones, A-GPS[3] (Assisted Global Positioning System) capability is now being included in new handset designs. A-GPS based positioning techniques allow the location of a phone – and hence, ordinarily, its user – to be determined indoors and in 'urban canyon' environments, where traditional standalone GPS is unusable. The scope of awareness can then be extended beyond the user's immediate proximity, allowing them to locate, for example, the nearest free expert in a particular subject at that moment.

Premium high-speed data services such as 3G allow smart phones to receive and transmit streaming audiovisual content. Again, there are opportunities to integrate this functionality with the Aware platform. A squawk box facility could be created, allowing unmediated conversation spaces to be constructed around an interest or topic. As numerical evaluations of interest and interest group relationships are accessible to the platform, a 'squelch' control could be included, which a user could adjust depending on how much less relevant background chatter they wished to hear. Speech recognition and synthesis could also be used, allowing, if the phone were paired with a Bluetooth earpiece, hands-free audio interaction with other Aware functions.

While Aware already supports administration and organisation of virtual meetings and presentations, a smart phone client could seamlessly integrate the means through which an individual remotely participates in or views an event – the phone 'ringing' to indicate that the event is about to start, just as it would if its user were invited to join a traditional conference call.

---

[3] www.globallocate.com

Inevitably, though, in a large organisation, a user will not be able to attend all events relevant to their interests. Speculative digital recording of television programmes that match a user's interest profile has become part of the cultural mainstream, first as a standalone hardware device and now, embedded within cable and satellite receivers. With Aware's strong emphasis on user's interests, it is natural to apply the idea to recordings of professional events. The ontology could be extended to allow individual segments of an event to be modelled, so that only relevant portions need be presented to the user. Local storage capacity of phones is likely to increase at a faster rate than the coverage of high speed mobile networks, at least in the short-to-medium future, so, even with a centrally accessible repository of recordings, it will be desirable to maintain a local library on the device, which can be accessed when the user is travelling. Local storage will, nonetheless, be finite and Aware provides a natural structure through which recordings can be prioritised: those of events that the user has attended or has marked in their diary can be considered most important, then recordings of other events advertised to the user and finally, those of any remaining events that are open to the user.

8.3 Supporting Research

Much of the supporting research that has informed this thesis is now several years old. Polanyi's famous work dates to 1966, Orr carried out his Xerox study in the late 1980s and, even in the three years that have elapsed since Nardi's investigation of intensional networks, commercial technologies available to support these networks have evolved rapidly. Notably, the roll-out of third generation (3G) mobile phone networks has significantly widened ownership of powerful mobile computing devices that offer a more credible and flexible platform for awareness, communication and community applications than PDAs or lab-constructed prototypes and opened new opportunities for continuous ambient video and audio connectivity. The nature of innovation-led work is also changing. While Western companies have long outsourced manufacturing capability to Asian countries, many are also now establishing offshore research and development centres and they will expect innovation support solutions to adapt to the different social and business cultures in these countries. There is no reason to believe that the basic academic principals identified in Chapter 2 will no longer hold true, but inevitably technical advances and increasing cultural diversity of the potential user base will hold new opportunities and challenges for innovation support systems and it is important that the academic research on which we rely keeps pace with them.

Finally, though the emphasis placed on awareness, community and communication reflects the importance of connectedness to innovation, we must concede that some of the world's most notable innovators have worked in glorious isolation. Further workplace analysis might reveal situations when *disconectedness* is beneficial – where a system like Aware should conceal or withold knowledge of relationships, not advertise them.

8.4 Conclusion

Innovation, essential to the success of organisations and industries, has largely been neglected by business software producers. Many a knowledge management tool has been developed to reinforce or retain existing practice and experience, but scant attention has been paid to the stimulation of new ideas. While existing information can be directly managed within computer systems, though, innovation can only be encouraged indirectly, by supporting a climate of free knowledge exchange between individuals. Recognising this, business theoreticians now advocate a community-based knowledge management strategy, which emphasises support for informal knowledge exchange networks. This strategy was shown to have three significant requisites: awareness, community and communication.

Awareness, community and communication have each been addressed many times individually, but the tools developed were handicapped by having access only to a very limited representation of the domain in which they operated. It was argued that individual interest, considered central to each of the three requisites, was especially poorly represented. In this thesis, for the first time, a single, carefully constructed conceptualisation, or ontology, forms the basis of a comprehensive suite of innovation support tools. Ontologies representing organisational structures have been published elsewhere, but, not only does the ontology presented here capture the reality of the innovation environment, rather than an idealised view for external consumption, it does so in a way that can be subjected to systematic evaluation. This fact is critical to the work presented here. It has allowed a platform to be created that can actively take decisions – *do things* for its users, rather than simply have them navigate around an information space.

This platform, Aware, has an active core, implemented using the JESS inferencing engine. Rules respond instantaneously to changes to its conceptual model, maintaining evaluations of relationships between topics, users' interests and interest groups. As the conceptual model is characterised by gradual change, the evaluation process is temporally distributed, so can be performed without impacting the performance of platform services. Because it is actively maintained and always available, the evaluation can viably be used by a wide range of different services. It provides the basis of a novel dynamically evaluated addressing mechanism for bulletins that allows users to *describe* who should see their message rather than have to list them by name. It also supports identification of individuals with particular interest or expertise – something that can, of course, be packaged as a user function with a dedicated visualisation, but can equally well be used to augment *any* function that returns a list of individuals. The platform's social network identification tool exemplifies the benefit gained by adopting a unified approach to innovation support, having an awareness both of friendships and past collaborations and of future opportunities to meet identified by reference to users' appointment diaries.

A client application has been developed as a component of a scenario of deployment within a single innovation-led technical organisation, but the ontology developed and the services built around it are flexible enough that they can support a wide range of innovation environments. Demonstrating this, a contrasting scenario has been explored, showing how Aware's services could benefit the organic farming community. Care has also been taken to minimise the burden that deployment of the platform would place on its users. While entirely capable of standalone operation, it has been recognised that this is not ideal when it requires that personal and organisational information already maintained elsewhere be duplicated locally. Provision has been made within the platform for it to be coupled with external sources of information by a lease mechanism. A development of existing semantic web triplestore infrastructure to support this has been proposed.

If implemented, proposed future developments, which include an intelligent scheduling system that would maximise the value gained from social gatherings and a suite of mobile phone tools that includes a potential killer-application for 3G networks, will further demonstrate the ability of this ontology-based platform to support organisational innovation networks.

References

Alani02 – Alani, H., O’Hara, K., Shadbolt, N., *ONTOCOPI: Methods and Tools for Identifying Communities of Practice,* in *Proceedings of 17th IFIP World Computer Congress,* Kluwer, 2002.

Alani03 – Alani, H., Dasmahapatra, S., O'Hara, K., Shadbolt, N., *Identifying Communities of Practice through Ontology Network Analysis*, in *IEEE Intelligent Systems*, 18(2):18–25, 2003.

Alavi01 – Alavi, M., Leidner, D.E., *Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues*, in *Management Information Systems Quarterly*, 25(1):107–136, 2001.

Allen83 – Allen., J.F., *Maintaining knowledge about temporal intervals,* in *Communications of the ACM*, 26(1):832–843, 1983.

Arundel02 – Arundel, A., Bordoy, C., *In-house versus ex-house: the sourcing of knowledge for innovation*, in De La Mothe, J., Link, A.N., *Networks, Alliances and Partnerships in the Innovation Process*, Kluwer, 2002.

Bellotti02 – Bellotti, V., Ducheneaut, N., Howard, M., Neuwirth, C., Smith, I., *Innovation in Extremis: Evolving an Application for the Critical Work of Email and Information Management,* in *Proceedings of the Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, June 25-28, 2002.

Benjamins98 – Benjamins, V.R., Fensel, D., Perez, A.G. *Knowledge Management through Ontologies*, in Reimer, U. *Practical Aspects of Knowledge Management PAKM 98. Proceedings of the Second International Conference in Basel*, Switzerland, October 1998.

Blackler95 – Blackler, F., *Knowledge, Knowledge Work and Organizations: An Overview and Interpretation*, in *Organization Studies*, 16(6):1021–1046, 1995.

Burgess98 – Burgess, C., Livesay, K., Lund, K., *Explorations in Context Space: Words, Sentences, Discourse*, in *Discourse Processes*, 25: 211–257, 1998.

Carotenuto99 – Carotenuto, L., Etienne, W., Fontaine, M., Friedman, J., Muller, M., Newberg, H., Simpson, M., Slusher, J., Stevenson, K., *CommunitySpace: Toward Flexible Support for Voluntary Knowledge Communities*, in *Proceedings of Changing Places Workshop*, London, April 1999.

Carter03 – Carter, S., Mankoff, J., Goddi, P., *hebb: An Infrastructure for Conveying Shared Interests*., UCB-CS-03-1259, Computer Science Division, UC Berkeley, July, 2003

Choo 98 - C. W. Choo, *The Knowing Organization*, Oxford University Press, 1998.

Clark96 – Clark, H., *Using Language*, Cambridge University Press, 1996.

Conway98 – Conway, S., Steward, F., *Mapping Innovation Networks*, in *International Journal of Innovation Management*, 2 (2): 165–196, 1998.

Cook 99 – Cook, S.,Brown, J., *Bridging epistemologies: The generative dance between organizational knowledge and organizational knowing*, in *Organization Science,* 10(4):381—400, 1999.

Cornejo03 – Cornejo, M., *Utility, value and Knowledge Communities*, online at www.kmadvantage.com/docs/km_articles/Utility_Value_and_K-Communities.pdf

Dahlbom95 – Dahlbom, B., Mathiassen, L., *Computers in Context: The Philosophy and Practice of Systems Design*, Blackwell, 1995.

Davenport98 – Davenport, T.H., Prusak, L., *Working Knowledge,* Harvard Business School Press, 1995.

Demarest97 – Demarest, M., *Knowledge Management – An Introduction,* online at www.noumenal.com/marc/km1.pdf

Domingue99 – Domingue, J., Motta, E., *A Knowledge-Based News Server Supporting Ontology-Driven Story Enrichment and Knowledge Retrieval,* submitted to the *11th European Workshop on Knowledge Acquistion, Modelling, and Management* (EKAW '99), 1999.

Drucker92 – Drucker, P.F., *Managing for the Future. The 1990s and Beyond*, Truman Talley, 1992.

Ducheneaut00 – Ducheneaut, N., Bellotti, V., *Email as habitat: An exploration of embedded personal information management*, in *ACM Interactions* 2001, September-October; 30-38, 2001.

Ducheneaut02 – Ducheneaut, N., Bellotti, V., *Ceci n'est pas un objet? Talking About Things in Email*, in *Journal of Human-Computer Interaction*, 18: 85–110, 2002.

Dumbill02 – Dumbill, E., *Finding Friends with XML and RDF*, online at www-128.ibm.com/developerworks/xml/library/x-foaf.html

Faber03 – Faber, A., Rood, T., Ros, J., *Evaluation of early processes in system innovation*, online at www.fu-berlin.de/ffu/akumwelt/bc2003/download/Faber_full.pdf

Fan01 – Fan, J., Barker, K., Porter, B., Clark, P., *Representing Roles And Purpose*, in *Proceedings of K-CAP'01*, Victoria, Canada, October 2001.

Firestone99 – Firestone, J.M., *Accelerated Innovation and KM Impact*, in *Financial Knowledge Management*, 1(1):54–60, 1999.

Fischer99 – Fischer, G., *Symmetry of Ignorance, Social Creativity and Meta-Design*, in *Proceedings of the third Creativity & Cognition Conference*, ACM Press, 1999.

Fischer01 – Fischer, G., Ostwald, J., *Knowledge Management: Problems, Promises, Realities and Challenges*, in *IEEE Intelligent Systems*, January/February 2001.

Fischer03 – Fischer, G., Ostwald, J., *Knowledge Communication in Design Communities*, in Bromme, R., Hesse, F., Spada, H., *Barriers and Biases in Computer-Mediated Knowledge Communication*, Kluwer, 2003.

Forgy82 – Forgy, C., *RETE: A fast algorithm for the many pattern/many object pattern match problem*, in *Artificial Intelligence*, 19:17–37, 1982.

Freeman00 – Freeman, L., *Visualizing Social Networks*, in *Journal of Social Structure*, 2000:1, 2000.

Genesereth87 – Genesereth, M.R., Nilsson, N.J. *Logical Foundations of Artificial Intelligence*. Morgan Kaufmann Publishers, 1987.

Geyer02 – Geyer, W., Cheng, L., *Facilitating Emerging Collaboration through Light-weight Information Sharing*, Conference Supplement of CSCW 2002. November 16-20, 2002. New Orleans, USA. 2002.

Goecks02 –Goecks, J., Cosley, D., *NuggetMine: Intelligent Groupware for Opportunistically Sharing Information Nuggets*, in *Proceedings of the ACM Conference on Intelligent User Interfaces*, ACM Press, 2002.

Glance98 – Glance, N., Arregui, D., Dardenne, M., *Knowledge Pump: Supporting the Flow and Use of Knowledge*, in Borghoff, U., Pareschi, R., *Information Technology for Knowledge Management*, Springer-Verlag, 1998.

Groth01 – Groth, K., *Using Social Networks for Knowledge Management*, online at http://www.ischool.washington.edu/mcdonald/ecscw03/papers/groth-ecscw03-ws.pdf

Gruber93 – Gruber, T.R., *A Translation Approach to Portable Ontologies*, in *Knowledge Acquisition*, 5(2):199—220, 1993.

Gruber95 – Gruber, T.R., *Towards Principles for the Design of Ontologies used for Knowledge Sharing*, in *International Journal of Human-Computer Studies*, 43(5/6), 907—928, 1995.

Guarino92 – Guarino, N., *Attributes and Arbitrary Relations*, in *Data and Knowledge Engineering*, 8, 1992.

Hammersley99 – Hammersley, M., Shadbolt, N. R., Golightly, D. A., Cottam, H. D., Riley, P. H., *Knowledge Engineering for Web Based Knowledge*

Management: A Case Study from Rolls-Royce plc., Knowledge Management, October 1999.

Hattori99 – Hattori, F., Ohguro, T., Yokoo, M., Matsubara, S., Yoshida, S., Socialware: Multiagent Systems for Supporting Network Communities, in *Communications of the ACM*, 42(3):55—61, 1999.

Heflin00 – Heflin, J., Hendler, J., *Searching the Web with SHOE*, in proceedings of *AAAI-2000 Workshop on AI for Web Search*. 2000.

Hildreth02 – The Duality of Knowledge, Information Research, Vol. 8 No. 1, October 2002

IDC96 – IDC. *Transition to the Information Highway Era*, in *1995-96 Information Industry and Technology Update.* 1996.

Jurisica99 – Jurisica, I, Mylopoulos, J., Yu, E. *Using Ontologies for Knowledge Management: A Computational Perspective*, in proceedings of *Annual Conference of the American Society for Information Science 1999, Washington, DC.* 1999.

Kalfoglou01 – Kalfoglou, Y., Domingue, J., Motta, E., Vargas-Vera, M., Shum, S.B., *MyPlanet: an ontology-driven Web-based personalised news service*, in *Proceedings of IJCAI'01 Workshop on Ontologies and Information Sharing*, 2001.

Kautz97 – Kautz, H., Selman, B., Shah, M., *The Hidden Web*, in *AI Magazine*, Summer 1997.

Koch00 – Koch, M., Lacher, M.S., *Integrating Community Services - A Common Infrastructure Proposal*, in *Proceedings of Knowledge-Based Intelligent Engineering Systems and Allied Technologies (KES2000), Brighton, UK*, 2000.

Koch03 – Koch, M., *Community Support in Universities – The Drehschreibe Project*, in Huysman, M., Wenger, E., Wulf, V., *Proceedings of International Conference on Communities and Technologies (C&T2003)*, Kluwer, 2003.

Kristoffersen99 – Kristoffersen, S., Ljungberg, F. *An Empirical Study of How People Establish Interaction: Implications for CSCW Session Management Models*, in *Proceedings of CHI 99*, 1—8, 1999.

Lambe02 – Lambe, P., The Autism of Knowledge Management, 2002. online at greenchameleon.com/thoughtpieces/autism.pdf

Lazega01 – Lazega, E., *The Collegial Phenomenon. The Social Mechanisms of Cooperation among Peers in a Corporate Law Partnership*. Oxford University Press, 2001.

Lenat90 – Lenat, D.B., Guha, R.V., *Building Large Knowledge-Based Systems: representation and inference in the Cyc project*. Addison-Wesley. 1990.

Liu02 – Liu, P., Curson, J., Dew, P., *Exploring RDF for Expertise Matching within an Organizational Memory*, in Banks Pidduck, A. (Ed.), *CAISE2002*, Springer-Verlag, 100–116, 2002.

Lindgren021 – Lindgren, R., *Competence Visualizer: A Knowledge Management System Generating Competence Patterns of Organizational Groups*, online at www.viktoria.se/results/result_files/167.pdf

Lindgren022 – Lindgren, R., Stenmark, D., *Designing Competance Systems: Towards interest-activated technology*, in *Scandinavian Journal of Information Systems* 14, 19–35, 2002.

Lindgren03 – Lindgren, R., Stenmark, D., Ljungberg, J. *Rethinking Competence Systems for Knowledge-Based Organizations*, in European Journal of Information Systems 12, 18–29, 2003.

MacDonald03 – MacDonald, J., Atkin W., Daugherity F., Fox, H., MacGillivray, A., Reeves-Lipscomb, D., Uthailertaroon, P. *Let's get more positive about the term 'lurker'*, online at www.cpsquare.org/edu/foundations

Majchrzak01 – Majchrzak, A., Ernst Neece, O., Cooper, L.P., *Knowledge Reuse for Innovation – The Missing Focus in Knowledge Management: Results of a Case Analysis at the Jet Propulsion Laboratory*, 2001.

Malhotra02 – Malhotra, Y., *Why Knowledge Management Systems Fail? Enablers and Constraints of Knowledge Management in Human Enterprises* . In Holsapple, C.W. (Ed.), *Handbook on Knowledge Management 1: Knowledge Matters*, Springer-Verlag, 577–599, 2002.

McArthur03 – McArthur, R., Bruza, P., *Discovery of Implicit and Explicit Connections Between People Using Email Utterance*, online at www.dstc.edu.au/Research/Projects/Infoeco/publications/discoveryscience03.pdf

McCarthy01 – McCarthy, J.F., Costa, T.J., Liongosari, E.S., *UniCast, OutCast & GroupCast: Three Steps Toward Ubiquitous, Peripheral Displays*, in *Proceedings of Ubicomp 2001*.

McCullough04 – McCullough, M., *Digital Ground: Architecture, Pervasive Computing and Environmental Knowing*, MIT Press, 2004.

McGuinness01 – McGuinness, D.L., *Ontologies Come of Age*, in Fensel, D., Hendler, J., Lieberman, H., Wahlster, W., *The Semantic Web: Why, What and How*, MIT Press, 2001.

Milestad02 – Milestad, R., Darnhofer, I., *Building Farm Resilience: The prospects and challenges of organic farming*, in *Journal of Sustainable Agriculture* 22(3): 81–97, 2002.

Millen02 – Millen, D., Muller, M., *Computer-Supported Communities of Practice*, online at http://www.usabilityviews.com/uv005533.html

Miller99 – Miller, W.L., Morris, L., *Fourth Generation R&D: Managing Knowledge, Technology and Innovation*, John Wiley & Sons, 1999.

Moller04 – Möller, K., Predoiu, L., Bachlechner, D., *Portal Ontology*, online at http://sw-portal.deri.org/papers/deliverables/D1_v1.2PortalOntology.pdf

Monk03 – Monk, A., *Common Ground in Electronically Mediated Communication*, in Carroll, J., *HCI Models, Theories, and Frameworks: Toward a Multidisciplinary Science*, Morgan Kaufmann, 2003.

Nabeth01 –Nabeth, T., Angehrn, A., Roda, C., *Towards Personalized, Socially Aware and Active Knowledge Management Systems*, CALT White Paper, 2001.

Nardi00 – Nardi, B., Whittaker, S., Bradner, E., *Interaction and Outeraction: Instant Messaging in Action*, in *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, ACM, 2000.

Nardi01 – Nardi, B., Whittaker, S., Isaacs, E., Creech, M., Johnson, J., Hainsworth, J., *Integrating Communication and Information Through ContactMap*, in *Communications of the ACM*, 45:4, 2002.

Nickols00 – *The Knowledge in Knowledge Management*, in *The Knowledge Management Yearbook 2000-2001*, Butterworth-Heinemann, 2000.

Nonaka00 – Nonaka, I., Toyama, R., Konno, N., *SECI, Ba and Leadership*, in *Long Range Planning*, 33(1): 5-34, 2000.

Oleary98 – O'Leary, D.E., *Enterprise Knowledge Management*, in *Computer*, March 1998.

Orchard01 – Orchard, R., *Fuzzy Reasoning in JESS: The FuzzyJ Toolkit and FuzzyJESS*, in *Proceedings of the ICEIS 2001, Third International Conference on Enterprise Information Systems*. 2001.

Orr96 – Orr, J., *Talking About Machines: An Ethnography of a Modern Job*, Cornell University Press, 1996.

Polanyi62 – Polanyi, M., *Knowledge And Being*, Routledge, 1962.

Polanyi66 – Polanyi, M., *The Tacit Dimension*, Doubleday, 1966.

Power98 - Power, Thomas M., *The Economic Pursuit of Quality*. M.E. Sharpe, Inc. 1988.

Scarbrough96 – Scarbrough, H., *The Management of Expertise*, St Martin's Press, London, 1996.

Scarbrough99 – Scarbrough, H., Swan, J., *Knowledge Management: A Literature Review*, Institute of Personnel and Development, 1999.

Scarbrough01 – Scarbrough, H., Swan, J., *Knowledge Communities and Innovation*, in Van Baalen, P., *Trends in Communication; Special Issue on Communities of Practice,* 7–20, Boom, 2001.

Schwartz92 – Schwartz, M. , Wood, D., *Discovering Shared Interests Among People Using Graph Analysis of Global Electronic Mail Traffic*, in *Communications of the ACM*, 36(8):78–89, 1992.

SeelyBrown95 – Seely Brown, J., Solomon Gray, E., *The People Are The Company*, in *Fast Company*, 1:78, 1995.

Seidler03 – Seidler-de Alwis, R., Hartmann, E., Gemunden, H.G., *The Role of Tacit Knowledge in Innovation Management*, presented at 19th Annual IMP Conference Lugano Schweiz, 4th – 6th September 2003.

Shadbolt04 – Shadbolt, N.R., Gibbins, N., Glaser, H., Harris, S., Schraefel, M.C., *CS AKTive Space or how we stopped worrying and learned to love the Semantic Web*, in *IEEE Intelligent Systems*, 19(3):41–47, 2004.

Sharp97 – Sharp, J., *Key Hypotheses in Supporting Communities of Practice*, online at www.tfriend.com/hypotheses.html

Solomon98 – Solomon, G.T., Winslow, K., Tarabishy., A., *The Role of Climate in Fostering Innovative Behaviour in Entrepreneurial SMEs*, online at www.usasbe.org/knowledge/proceedings/1998/29-Solomon.PDF

Skyrme99 –Skyrme, D., *The Promise and Challenge of Knowledge Markets*, in Davenport, T. H., Prusak, L. *Working Knowledge*, Harvard Business School Press, 2000.

Stenmark02 – Stenmark, D., I*nformation vs. Knowledge: The Role of Intranets in Knowledge Management*, in *Proceedings of the 35th Hawaii International Conference on System Sciences*, 2002.

Swan99 – Swan, J., Robertson, M., Newell, S., *Knowledge Management: When will People Enter the Debate*, in T.K.Kkl (ed.) Proceedings of the 22nd Information Systems Research Seminar in Scandinavia (IRIS 22), 1999.

Swan00 – Swan, J., Newell, S., Robertson, M., *Limits of IT-based Knowledge Management Initiatives for Interactive Innovation Processes: Towards a Community-Based Approach*, in *Proceedings of the 33rd Hawaii International Conference on System Sciences*, 2000.

Swan02 – Swan, J., Scarbrough, H., Robertson, M., *The Construction of Communities of Practice in the Management of Innovation*, online at http://www.ki-network.org/downloads/ikon_papers/community.pdf.

Tsoukas96 – Tsoukas, H., *The Firm as a Distributed Knowledge System: A Constructionist Approach*, Strategic Management Journal, 17, Winter Special Issue, pp. 11–25. 1996.

Tuomi99 – Tuomi, I., *Data is More Than Knowledge*, in *Journal of Management Information Systems*, 16(3):103-118, 1999.

UNESCO03 – *The Culture of Innovation and the Building of Knowledge Societies*, UNESCO Buereau of Strategic Planning, 2003.

Uschold96 – Uschold, M. Gruninger, M. *Ontologies: Principles, methods and applications*, in *Knowledge Engineering Review*, 11(2), 1996.

Venters02 – Venters, W., *Literature Review for C-Sand: Knowledge Management*, online at www.c-sand.org.uk/Documents/WP1001-02-KMLitRev.pdf

Veryzer98 – Veryzer, Robert W. *Discontinuous Innovation and the New Product Development Process*, in *Journal of Product Innovation Management* **15** (4). 1998.

Vogiazou03 – Vogiazou, I.T., Dzbor, M., Komzak, J., Eisenstadt, M., *BuddySpace: Large-Scale Presence for Communities at Work at Play*, Submitted to Inhabiting Virtual Places, Workshop, E-CSCW-03, September 2003, Helsinki, Finland. 2003.

Want02 – Want, R., Pering, T., Danneels, G., Kumar, M., Sundar, M., Light, J., *The Personal Server - Changing the Way We Think about Ubiquitous Computing*, in *Proceedings of Ubicomp 2002: 4th International*

Conference on Ubiquitous Computing, Springer LNCS 2498, Goteborg, Sweden, Sept 30th-Oct 2nd, 2002, pp194–209. 2002.

Weskamp04 – Weskamp, M., *Social Circles. Mailing List Social Visualisation*, online at www.marumushi.com/apps/socialcircles/, 2004.

Whitley00 - Whitley, E.A., *Tacit and Explicit Knowledge: Conceptual Confusion around the Commodification of Knowledge*, in *Knowledge Management: Concepts and Controversies*. BPRC, Warwick University, 2000.

Wilson02 – Wilson, T.D., *The Nonsense of Knowledge Management*, in *Information Research*, 8:1, October 2002.

Zdrahal01 – Zdrahal, Z. *Online Workshop: The Role of Ontologies in Knowledge Management* (online at www.knowledgeboard.com/item/50466)