Awareness of Organizational Expertise

Mark Maybury[2], Ray D'Amore[1], and David House[1]
The MITRE Corporation
[2]202 Burlington Road
Bedford, MA 01730
[1]1820 Dolley Madison Blvd
McLean, VA 22102

{maybury, rdamore, dhouse}@mitre.org

Abstract

This article describes automated tools for increasing organizational awareness within a global enterprise. The MITRE Corporation is the context for the current work, however the tools and techniques are general and should apply to a wide variety of distributed, heterogeneous organizations. These tools provide awareness of team members and materials in virtual collaboration environments as well as support for automated discovery of distributed experts. The results are embodied in three systems: MITRE's Collaborative Virtual Workspace (CVW), Expert Finder and XpertNet. CVW is a place-based collaboration environment that enables team members to find one another and work together. Expert Finder is an expert skill finder that exploits the intellectual products created within an organization to support automated expertise identification. XpertNet addresses the problem of detecting extant or emerging classes of expertise without a priori knowledge of their existence. Both Expert Finder and XpertNet combine to detect and track experts and expert communities within a complex work environment. After describing the background of knowledge management at MITRE, this article describes the architecture and use of collaboration and expert finder systems to enhance organizational awareness, provides some principles of expertise, and concludes with an outline of future research directions.

1. A Framework for Awareness

Awareness encompasses knowledge of people, their presence and availability, their activities, the information they produce, and tools they use. Awareness can include static information (e.g., the physical location of a specialty department in a university) as well as more dynamic information (e.g., the location of a user, the state of the weather, or the price of a stock). Awareness may involve explicit information (e.g., the logon time of a user) or implicit information (e.g., the user's interests as indicated by their browsing behavior).

People. Awareness of people can encompass typically static properties such as their name and gender as well as more dynamic ones such as their time of login, current physical or virtual location, current title, or availability (e.g. on or off-line, busy). In the context of conferencing, this might include their role in a session (e.g., chairperson, scribe, facilitator, member) as well as properties such as permissions associated with information and/or tools. Users may have specific roles within particular work contexts (e.g., creating expert knowledge, applying it, or teaching it). Moreover, individuals have particular expertise, skills, and explicit knowledge that is key to supporting effective matchmaking. A special case involves software agents that might either act as proxies or avatars for people or may be first-order participants in a session, in which case they too may have additional properties such as location or permissions.

Information. Awareness of information includes properties such as the type of information (e.g., image, document, briefing), the content of the information (e.g. who, what, when, where, how), status (e.g., draft, final), security properties, its location, and current availability. Users awareness regarding information can be managed by alerting folks about particular changes in event properties (e.g., initial publication, change of access properties, moving from draft to final) based on some model of their interest in information objects. Other kinds of awareness could include awareness of the evolutionary and/or access history of the document.

Tools. Awareness of tools might encompass their availability, state, and/or current people using them. Of particular interest is the ability to have awareness in both the use of tools synchronously and asynchronously. For example, users of a simultaneous whiteboard could see others; however, attributing their annotations during a session with the particular users that create them and storing these provides subsequent users with awareness of the original creators and/or the evolution of the product. Similarly, users may benefit from awareness of the state of an application (e.g., open, closed) or its event history.

Activities. Awareness of activities includes awareness of project tasks, information flow and work flow within and across organizations. This type of awareness concerns when tasks are performed, with whom they are performed, and how they are performed.

We have explored awareness in a variety of applications including awareness in the context of enterprise knowledge management and collaboration, as well as awareness of experts in order to support problem solving or team creation. The remainder of this paper addresses the role of awareness in knowledge management, prior research in this area, a description and evaluation of MITRE's CVW, Expert Finder and XpertNet prototypes, principles of expertise, and plans for future research.

2. Awareness to Support Knowledge Management

The MITRE Corporation serves as an objective, non-profit corporation whose mission is to serve the public interest by creating solutions to pervasive, cross-organizational problems facing government in civil aviation, tax administration, and national security. MITRE's mission is to enable innovation, integration, and collaboration within and across public sector agencies. Awareness and expertise management are core functions in achieving this objective. The corporation's sponsors need help in being agile in response to world changes, ensure they are interoperable with national and international partners, and counter the loss of government experts. MITRE has approximately 4,500 people are distributed globally (in the US, Europe, and Far East) and include both technical and mission or operational experts. Regular interaction among technical and domain experts distributed throughout headquarter and sponsor collocated units enables rapid and high quality creation of solutions that make a difference.

Managing and enabling MITRE's distributed resources is a daunting challenge. Our strategy, as illustrated in Figure 1, is to provide full knowledge life-cycle management, to include Web-based expert finding services, capture/reuse of knowledge and lessons learned, project and partnership creation (leader identification, team formulation, team facilitation/collaboration), and knowledge delivery. As is indicated in the figure, awareness, of individuals, events and information is a central notion to the success of our knowledge management strategy.

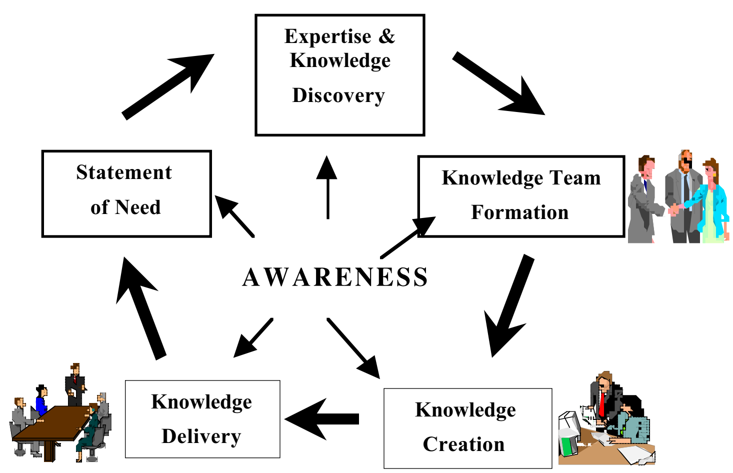


Figure 1: Knowledge Management Processes

To enable effective distributed knowledge management, MITRE employs global video teleconferencing, an innovative MITRE Information Infrastructure (MII) intranet, as well as public key infrastructure (PKI) enabled extranet services. Knowledge resources such as "ask the expert," on-line knowledge repositories of risk management experience (RAMP), and system engineering lessons learned (Systems Engineering Program Office) are also available.

Tools for staff and project discovery, information sharing (e.g., transfer folders) and virtual place-based collaboration such as MITRE's CVW (cvw.mitre.org) are used to enable agile team formation and support. Finally, a collaborative culture, organizational structures (e.g., technical councils, technical centers of excellence), special roles (e.g., sponsor focal points, strategic outcome leaders) as well as regular knowledge exchange meetings (called Technical Exchange Meetings or TEMs) support the process shown in Figure 1.

Effective expert and expert team discovery and support raise a series of complex issues. These include what is an expert? How can various levels of expertise or competence be characterized (e.g., student, master)? How can experts be authenticated or validated? How does and individual's level of expertise evolve over time? How can this be systematically supported in a cost-effective manner?

3. Prior Research

With the rapid growth of synchronous messaging, a number of commercial tools have emerged to support real-time on-line communication such as AOL Instant Messenger, Lotus SameTime and Microsoft NetMeeting. In the context of conferencing services, these tools provide buddy lists and other forms of user presence information. Notable is the Internet Engineering Task Force (IETF) effort to establish an Instant Messaging and Presence Protocol (IMPP) to establish a uniform manner of sharing information across applications about individual identity, communication capabilities and availability.

Before individuals can converse, however, they need to know who to talk with. Several investigators have explored automating the process of expertise discovery and assessment. For example, in the Dataware II Knowledge Directory[1], experts can self-nominate and subsequently be discovered through directory search, however, this manual process is expensive to maintain and becomes quickly out of date. In contrast, Autonomy[2] analyzes users' search and publication histories to determine concepts that are indicative of their expertise. Yenta (Foner, 1997) and Tacit KnowledgeMail[3] determine user expertise from email message traffic. MIT's ExpertFinder (Vivacqua 1999) instruments software library usage to determine expertise level. Referral Web from AT&T (Kautz et al. 1997) provides access to experts across an enterprise or community, aiming to make the basis for referral transparent to the user. It generates social networks based on bibliographic information and supporting context to deduce actual experts and associated referral paths. U.S. West's Expert-Expert Locator (Streeter & Lochbaum 1988) also finds experts across an enterprise.

Abuzz's Beehive[4] is one of many systems that provide an on-line community environment to support question/answer dialogues between users and registered "experts." Users can learn from other user's question/answer dialogues posted under specific topics such as *cooking*. Communities of experts are grouped in *web circles* that provide a domain specific context for registering as an expert, for users to ask questions or initiate a group discussion. This is similar to The Answer Garden (Ackerman and Malone, 1990) which categorized questions into ontology, which could be browsed by users to find questions/answers similar to their own question. If users did not find a related question they were referred to an expert. The emerging on-line commercial systems attempt to also track each experts' performance; and the general trend is to use user ratings and experts response times as a basis for measuring competence.

Essentially, social filtering is used to qualify the level of expertise of registered experts. As such systems often suffer from the cold-start problem where there is a mismatch between the number of experts and users. In some cases experts outnumber users, discouraging experts' participation or affecting revenue. In other cases, there is a dearth of experts (or qualified experts) and users become frustrated because of poor response times or low quality answers. While these systems (e.g., XperSite.Com[5]) present interesting expertise management paradigms a number

---

[1] Dataware Knowledge Management Systems White Paper *(http://www1.dataware.com/forum/kms/kmsfull.htm)*

[2] Autonomy Technology White Paper *(http://www.autonomy.com/tech/wp.html)*

[3] Tacit Knowledge Systems' KnowledgeMail (*http://www.tacit.com*)

[4] Abuzz "Ask Anything" (*http://www.abuzz.com*)

[5] XperSite.com (*http://www.xpersite.com/*)

of core problems remain. including representing and measuring an expert's qualifications, as well as matching questions to the appropriate experts.

MITRE's approach is unique by combining a dynamic search/indexing system with a set of community-of-practice discovery and tracking tools. This approach also differs from previous systems by offering named entity extraction to find experts mentioned in unstructured text.

Subsequent sections describe MITRE's applied research. This encompasses a collaborative virtual workplace, expert finders, and expert community finders.

4. Expert Finding

Collaborative environments are only effective when experts are accessible within them and those experts are able and willing to share their knowledge. In the remaining sections of this article we describe tools which address some of the following fundamental aspects of expertise management:

- *Expertise Identification:* There are a wide variety of digital sources (e.g., published document, tool logs) that can be exploited to determine who in an organization is associated with particular domains of expertise and what degree or level of expertise is associated with them. For example, in MITRE's environment, some of these digital sources have been experimentally shown to contribute more to this task than others, and a combination of sources tends to determine the best results.

- *Expertise Classification*: By assessing multiple sources of evidence, it is possible to characterize the type and level of expertise of an individual and then group individuals into classes of experts. For example, as described below, MITRE's tools create specific expertise classes using statistical clustering techniques and social network analysis.

- *Expertise Validation*: Identified experts need to be qualified as actual experts. Several methods include human assessment, making reference to evidence of their expertise (e.g., qualifications, resume, publications), or automated user feedback mechanisms which capture positive and negative feedback from interaction with those experts. Privacy concerns may limit the degree to which measurements of expert performance are shareable.
- *Expertise Collaboration*: In our increasingly complex world, solutions to problems often require either communities of experts or diverse ranges of expertise that need to be brought together to solve the complex problems.
- *Expertise Monitoring*: Expertise continuously changes and requires awareness of this dynamic.
- *Expertise Access*: True expertise is rare and expensive. Often access is controlled, either informally or formally, either by the expert themselves or their management.

Each of these aspects are concretely illustrated in the descriptions below of Expert Finder and XpertNet. First, however, we describe a virtual collaboration environment which supports access to and collaboration with and among experts.

5. Place-based Virtual Teams

MITRE's Collaborative Virtual Workspace (CVW) (Spellman et al. 1997) is a place-based collaboration environment based on a "virtual building" metaphor where teams can communicate, collaborate, and share information, regardless of their geographic location. CVW takes virtual meetings one step further and enables virtual co-location through persistent virtual rooms, each incorporating people, information, and tools appropriate to a task, operation, or service.

Awareness of people and materials is illustrated in Figure 2 showing screen shots of CVW. Awareness of other users, information and tools pervades the interface. For example, in the CVW Client Interface in the upper left-hand side of Figure 2, the interface displays text chat, users in the current room, and shared data. The virtual building floor plan shown in the middle upper portion of Figure 2 gives the users a sense of where they are in virtual space.

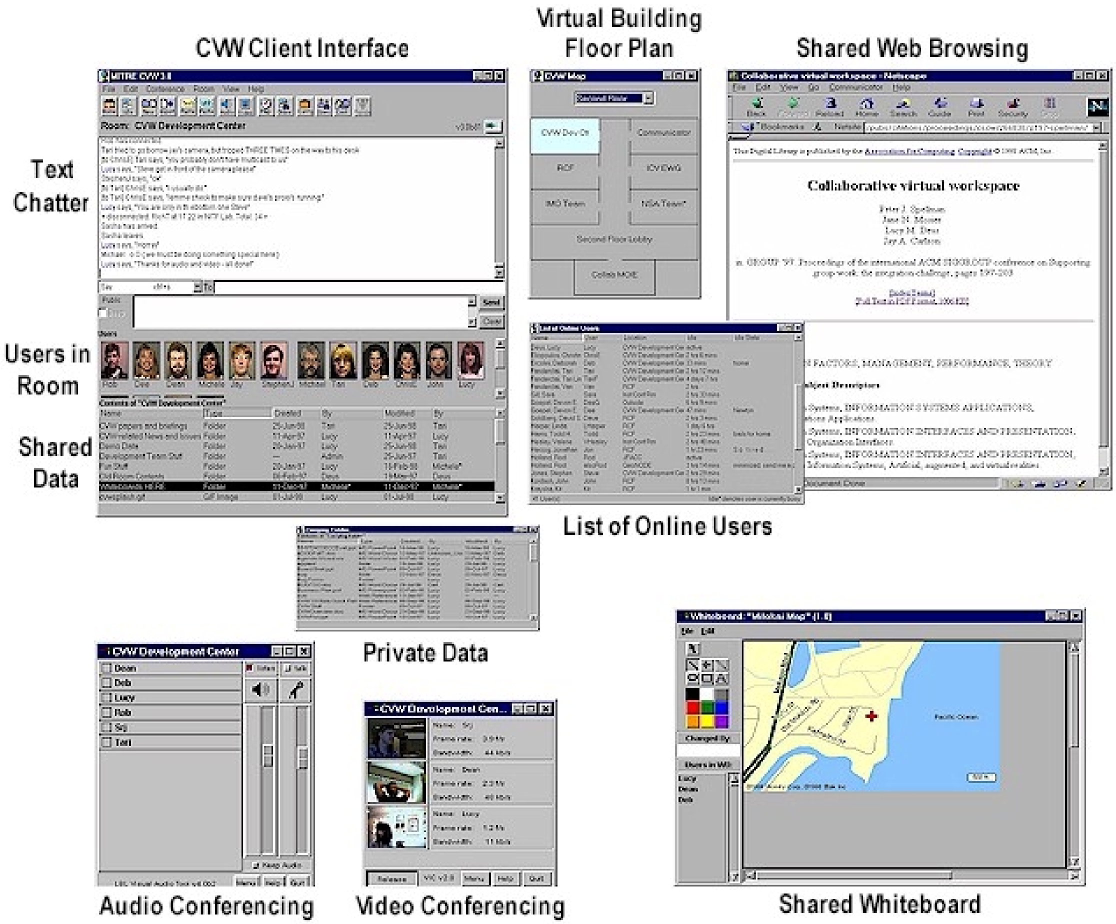


Figure 2: CVW

Navigational facilities, such as the ability to browse rooms and floors just as one could do in a physical building or the ability to move immediately to a virtual room or user by typing in that identifier, enhances the awareness of the user's environment. Moreover, in each interaction device such as chat, whiteboard, and audio/video conferencing, active participants are displayed and individual contributions are visibly attributed to individuals. For example, in the audio conferencing tool displayed in the lower left-hand portion of Figure 2, the current speakers in the teleconference are listed in the left-hand side of the display. As each speaks, their name is highlighted.

Similarly, in the shared whiteboard displayed in the lower right hand side of Figure 2, the active users in the whiteboard conference are listed in the left hand side of the display just below the color palette. As users perform white board actions, their identities are signified in the cursor via an abbreviation and their persistent annotations are attributed with their identity. This allows users who join in subsequently in the session to have awareness of the source of annotations to the whiteboard.

6. MITRE's MII Expert Finder

Distribution of staff, decreasing project size, and cost/time pressure are driving a need to leverage enterprise expertise by quickly discovering who knows what and forming expert teams. Those in need typically have little or no means of finding experts other than by recommendation. Unfortunately, busy experts do not have time to maintain adequate descriptions of their continuously changing specialized skills. Past experience with "skills" databases at MITRE indicates that they are difficult to maintain and quickly outdated, especially considering an industry-wide average employee attrition rate of greater than 10% in high-tech fields.

MITRE's MII Expert Finder (Mattox, Smith, and Seligman 1998) fills this gap by mining information and activities on the MII related to experts and providing this in an intuitive fashion to end users. Figure 3 illustrates the system in action. In this example, a user is trying to find human computer interaction experts in the corporation. When the user searches using the term "human computer interaction," the system ranks employees by the number of mentions of a term or phrase and its statistical association with the employee name either in corporate communications (e.g., newsletters) or based on what they have published in their resume or document folder (a shared, indexed information space). Integrated with MITRE's corporate employee database, employees are ranked by frequency of mentions, pointing to sources in which they appear.

Figure 3: MII Expert Finder "Human Computer Interaction" Example

The MII Expert Finder exploits the MII, and thus avoids maintaining information internally. By doing so, Expert Finder operates in real-time, using the most recent information available to locate experts. The MII contains many different sources of information that can be used to locate relevant expertise. Staff members can easily and quickly publish documents in individual staff "document" folders on the MII. These include technical papers, presentations, resumes, and home pages. Also, information is published about MITRE employees in project descriptions, announcements, and internal and external newsletters. At MITRE, a common text search engine indexes all of these documents.

6.1 MII Expert Finder Process

The MII Expert Finder works by integrating expertise-related information found through queries to the MII's search engine. In and of itself, each source of employee information mentioned above is generally not sufficient to determine if an employee is an "expert" in a particular topic. Expert Finder relies on the combination of evidence from many sources, and considers someone an expert in a particular topic if they are linked to a wide range of documents and/or a large number of documents about that topic.

Expert Finder has two main components: a set of scripts to call the MII's search engines and process the results, and a Web-based user interface. The steps that Expert Finder goes through when processing a query are shown in Figure 4. The user is initially presented with a Web-based CGI form that contains a search box for entering the query. Because some queries with many documents can take 30-45 seconds to process, the results of some commonly asked queries are "cached" (saved) by the system, and the user is offered the option of viewing these cached results when available. If the resulting query has not been cached, the MII Expert Finder takes a keyword phrase (e.g. "chemical weapons") and passes it to the underlying search engine which then returns a set of documents (including resumes, home pages, presentations, newsletters, etc.) as a set of hyperlink pointers.

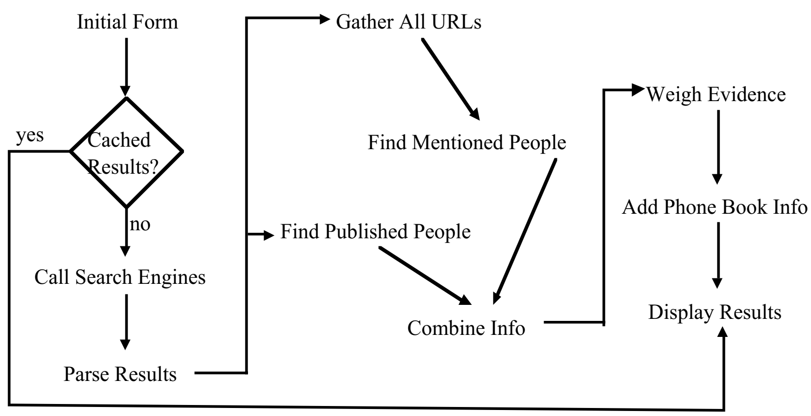


Figure 4: MII Expert Finder Process

As described above, the documents used by the MII Expert Finder fall into two main categories: documents about a topic that are published by an employee and documents that mention employees in conjunction with a particular topic. For the self-published case, the MII Expert Finder relies on the number of documents published by an employee about a given topic to provide an "expert score" for that employee. The only exception is that of an employee's resume, which is given additional weight as a self-definition of an individual's expertise.

Managing the second class of documents, those that mention employees and topics, is more complicated. Whereas with the self-published documents there is an explicit linkage between the documents and employee (they are indexed by employee number), with documents that mention employees, this linkage must be derived from the underlying text. The first step after the documents have been returned from the search engine is to locate the proper names within the text using a commercial product that tags names within a document[6]. Once the names have been located, the next step is to associate them with the query topic. All the documents returned by the search engine contain the query string, but there are several distinct types of documents, and each type has a structure that must be exploited differently. For example, MITRE publishes an internal newsletter that contains short paragraphs describing accomplishments by MITRE staff (e.g., "Dr. John Smith presented a paper titled 'Intelligent Agents and Data Management' at the Tenth International Conference on Autonomous Agents"). In this case, the MII Expert Finder uses proximity as a metric for expertise (i.e., the name "John Smith" is five words away from the keyword phrase "Intelligent Agents"). In addition, the HTML tags of the document can be used to further refine the score. In this case, a paragraph marker (<p>) breaks the linkage between a name and a topic since the announcements are always contained within a single paragraph. However, the heuristics used for the newsletter do not work well with other documents. For example, in technical reports the author's names appear at the top of the document and may be several paragraphs away from the relevant keywords and therefore require different heuristics to determine the linkage between names and keywords.

Once each document has been examined, the evidence gathered about each person found is combined into a single score for that person. That person's name is then matched against a database of MITRE employees, ordered by their final score and displayed in a 2-framed Web browser window. This means that the people who are most likely to be experts are displayed at the top of the list (see example in Figure 3). In the left frame of the results window is a photo of the employee, the employee's name, and the employee's department (all taken from the MII phonebook). If the user clicks on an employee's photo or an employee's name, then the right frame displays detailed employee contact information (including name, telephone number, email address, department, room number, etc.) taken from the MII phonebook, followed by an ordered listing of hyperlinks to the relevant documents that are associated with an employee's expertise.

6.2 MII Expert Finder Performance

Overall, the results obtained by the MII Expert Finder system are quite good. The original goal was to place a user within one phone call of an expert. That is, even if the persons listed as the result of an MII Expert Finder query were not the experts, they would be able to provide the name of someone who was. However, in the majority of the

[6] NameTag from IsoQuest Corporation.

cases tested, reasonable candidates for the title “expert” are listed as the top three or four candidates, where the likelihood of randomly selecting a correct expert is the total number of experts divided by 4,500 total staff, often significantly less than a one percent chance of getting any right. Table 1 illustrates preliminary results contrasting the performance of ten technical human resource managers, professionals at finding experts, with the MII Expert Finder for the task of identifying the top five corporate experts in specialty areas listed in the table. The first column in Table 1 shows the degree of inter-subject variability in reporting experts (measuring percentage of agreement of first, second, and third of five experts). Columns 2 and 3 show results for *precision*, the degree to which a staff member found by the MII Expert Finder is considered expert by humans, and *recall*, the degree to which a priori human-designated experts are found by the MII Expert Finder. Table 1 uses harsh measures where precision measures how many of the top-five MII Expert Finder results were also identified as expert by humans. In contrast, recall measures how many of the top-five experts that humans identified were included in the top five MII Expert Finder results. In spite of human variability (e.g., note the difficulty humans had in identifying chemical and network security experts), the MII Expert Finder works remarkably well except in network security and collaboration (ironically a result of few expert collaboration staff publishing on the MII, perhaps because they use specialized collaboration environments that were not instrumented).

Expert Area	**Human Agreement (1st, 2nd, 3rd)**	**ExpFinder Precision**	**ExpFinder Recall**
Data mining	70%, 49%, 24.5%	60%	40%
Chemical	40%, 8%, 0.8%	60%	40%
HCI	90%, 36%, 11%	60%	40%
Network Security	50%, 10%, 0.4%	20%	20%
Collaboration	70%, 35%, 17.5%	5%	5%
AVERAGE	**63%, 28%, 11%**	**41%**	**29%**

Table 1. Human and Expert Finder Performance

7. MII XpertNet

MITRE Technology Centers conduct applied research in a number of technology areas related to the sponsor's mission. As such they often partner directly with project departments; forming teams with diverse but complementary skills and problem knowledge. Organizationally, staff working related technologies and problems could be modeled as social networks that form the basis for abstracting expertise. According to [Ackerman et al., 1999], expertise networks can be defined as "… specializations of an organization's social network. They consider not only how people are socially arranged but what expertise they have and trade." Mapping expertise facilitates collaboration and provides a basis for directing strategic investments and optimizing resource allocation. XpertNet is designed to extract expertise networks and integrate them into the overall expertise management system. XpertNet works without user queries to identify expertise areas; a distinction between it and other expertise locator tools.

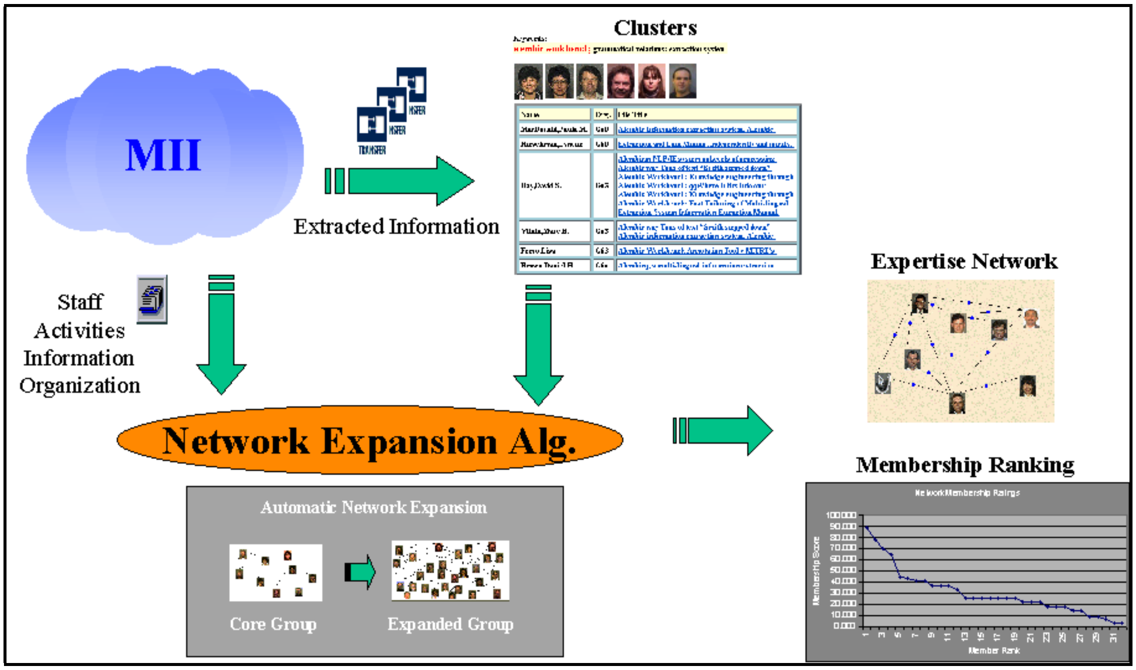


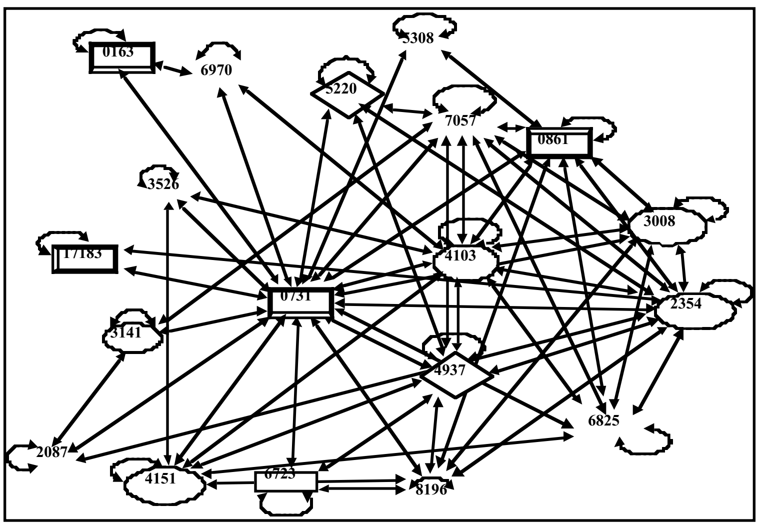
Figure 5: XpertNet Network Generation Algorithm

Figure 5 illustrates the overall data and control flow of XpertNet. XpertNet uses statistical clustering techniques and social network analysis to glean networks or affinity groups consisting of people having related skills and interests. Networks are extracted from various work contexts or activities such as projects, publications, and technical exchanges. Clusters are mapped to an expertise area description, a membership list consisting of MITRE technical staff and their degree of membership, and a list of content items on which the cluster is based. Information from published documents, public share folders, project information, and other sources are used to assess level of expertise. Higher levels of expertise are associated with factors such as document authorship, explicit reference or citation, network centrality, personal Web pages, and project membership. Lower expertise levels reflect fewer expertise indicators and possibly counter-indications such as being a member of the administrative staff. Currently, XpertNet incorporates domain independent models of expertise; domain-specific expertise models are expected in niche technology areas (e.g., Perl programming).

XpertNet can be directed at specific MII workspaces. Generally, items are simply selected from each workspace; however, large workspaces containing thousands of items are sampled to reduce processing costs. For example MITRE Public Share Folders consist of a wide range of documents provided in a user-controlled directory for sharing with interested MITRE staff. Share folders often contain hundreds of items or more, are quite diverse and may contain items that are fairly recent or several years old. Given the wide disparity in folder characteristics, a sampling frame ensures that a representative number of items are extracted for each user, that topics are adequately covered, and that the sample reflects recent work. In effect, the sampling scheme provides a statistically sound basis for working with a subset of a potentially large population of work artifacts. A signature is generated for each object sampled. Each signature contains person identifiers (e.g., employee ID, phone number, position level), workspace descriptor (e.g., Technical Exchange Meeting, Research Project, and Public Share Folder), and expertise indicators (e.g., descriptive phrases) extracted from activities or stored content.

Extracted signatures are input to a clustering algorithm that generates core topics consisting of a set of related objects. The clustering algorithm uses a dynamic merging scheme to collapse lower-level clusters into larger groupings-- the core cluster; the level of merging can be controlled consistent with the topic specificity selected for the system. It is possible to represent general skill areas that capture variations of, e.g., data mining expertise, by merging these into one large core cluster -- sub-network. A network expansion algorithm "grows" the sub-network (into an XpertNet) by exploiting several types of linkages: content links relating two or more members, social network links (e.g., project co-membership), and organizational links (e.g., members of the same department). Content links form the core XpertNet. Social network links are obtained from projects and other workspaces that have a formal membership and have a work context similar to the cluster core. Merging the content-based links with the membership-based links forms a consensus network. Organizational links are used to discern the organizational distribution of the experts; this has value when forming project teams or mapping work centers. Experts are ranked based on their network "significance" to include the summed relevance of content items, the number of links (in-links and out -links), and the membership weight in each project (or workspace). In essence, a person who has strong representation on relevant projects or activities, has a number of work artifacts associated with the expertise area, and who is highly connected to related experts is central to the expert network. In some cases, members that have complementary skills or serve in supporting roles are incorporated into the expert network. An example of an expertise network, with members identities masked out, is provided in Figure 6. In this "map," nodes represent people within the organization that are "involved" in our natural language processing work. The geometric shape of each node relates to the particular organizational entity (division, department, etc.) of each network member. In general, the thickness of the links corresponds to the strength of the association between these entities.

Figure 6: An Expertise Network



Expertise indicators are used to generate expertise ratings for each network member, as illustrated in Figure 7. Ratings are used to rank experts for selection or to highlight experts within a network visualization.

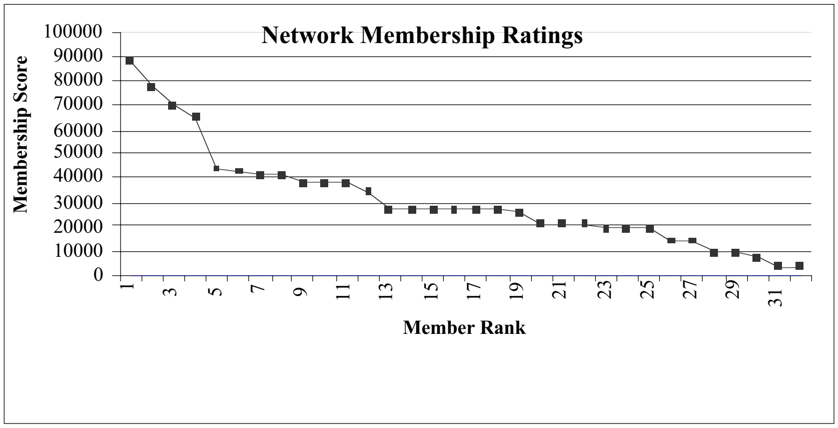


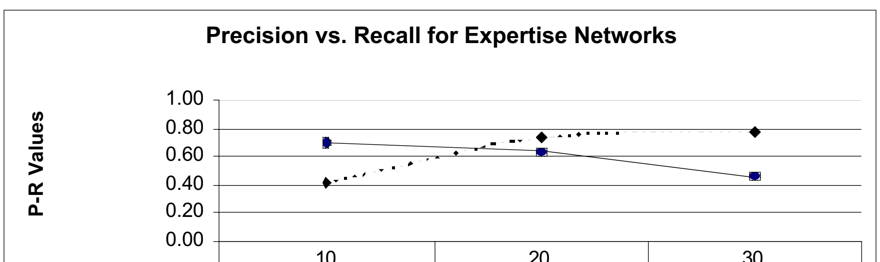
Figure 7: Expertise Ratings

7.1 XpertNet Evaluation

The current XpertNet algorithm is under evaluation at MITRE's Information Technology Center. Initially, expertise networks, generated from user surveys, were compared to automatically generated expertise networks. The survey was administered to roughly 10% of the Center's technical staff. Each survey respondent nominated MITRE staff that had expertise in the specified technology area. Respondents nominated those they worked with as well as other known corporate experts, typically outside their home organization. Initial results yielded four core expertise areas: Collaboration, Knowledge Management, Advanced Instructional Training, and Language Processing. These areas match up fairly well with 4 separate departments within the Center; however, many of the identified experts come from other organizations outside the Center.

Standard precision and recall measures, traditionally used in information retrieval experiments, were used to assess the overlap between the manual and automatically generated networks. The surveys did not provide a ranking of experts, so precision at a particular cutoff was defined as the percentage of manually identified experts who were in the automatically generated list. Similarly, recall at a specified cutoff is the percentage of automatically ranked members who were in the total manually generated expert list. As shown in the combined table/graph in Figure 8, approximately 70% of the top ten automatically identified experts were in the manually identified list. Precision dropped about 10% when computed over the top 20. Recall, at the top 10 cutoff, was also fairly high, but this is partly a function of the relatively small number of experts identified by humans. Looking at the top 30, approximately 75% of the experts were identified automatically with approximately 50% accuracy.

Figure 8: Average Precision vs. Recall for Expertise Networks



	10	20	30
Precision	0.70	0.63	0.46
Recall	0.42	0.73	0.77

Cutoff Values

A pilot system is being developed to automatically extract communities of practice and capture areas of expertise within each community identified. As shown in Figure 9, the system generates a "community card" for each community of practice located. Each community of practice is summarized at the top level in terms of the most central community members (i.e., those with the highest centrality), the organizational unit in which most of the work is centered, key projects that support the work when available, and a community name consisting of a set of descriptive phrases automatically extracted from "work traces" located in the MITRE environment. Each community card can be opened to reveal additional detail to include a ranked community membership list, a community graph visualization showing which members are tied together based on overlapping projects or organizational assignments and a list of work objects used to identify the work focus of the community. The work objects (or traces) include documents, web pages, project descriptions, labor charges, and other organizational information. The current results suggest the feasibility of automatically extracting expertise networks. The system provides a basis for identifying experts, staff with complementary skills, administrators that may act as information hubs, and large organizational entities that support the specified work area. In addition, it provides an evolving (and potentially observable) record of the changes and shifts in both technical and organizational focus within our Technology Center. However, there are a number of research issues, discussed below, to be addressed in the next generation of XpertNet. First, however, we articulate principles extracted from prototyping and piloting the above technologies.