Journal of Universal Computer Science, vol. 8, no. 5 (2002), 491-505
submitted: 4/2/02, accepted: 3/5/02, appeared: 28/6/02 © J.UCS

Knowledge on Demand: Knowledge and Expert Discovery

Mark T. Maybury
(The MITRE Corporation
202 Burlington Road
Bedford, MA 01730
maybury@mitre.org
www.mitre.org/resources/centers/it)

Abstract: This article outlines new technologies in the areas of automated expertise finding, expert network discover, virtual place-based collaboration, and automated question answering. We illustrate each of these areas with implemented and in some cases empirically evaluated systems. Collectively, these illustrate new methods for automatic discovery of knowledge, experts, and communities in an effective and efficient manner.

Keywords: knowledge management, knowledge acquisition, natural language, distributed collaboration
Categories: H, H.1.2, H.3.3, H.5.1, H.5.2, S.D I.2.1

1 Introduction

Knowledge creation is accelerating, driving an increased need for more effective management of knowledge [Morey et al. 2000]. For example, in the US there are more than 300,000 new patent applications annually which result in approximately 160,000 new patents added to the more than 6 million current patents. Whereas the size of the library of congress is 33 terabytes (growing at about 7,000 materials a day), one estimate is that the long distance communications in the U.S. alone in 1999 were 70,000 terrabytes. Digital internet transaction surpassed telephone communications volume in the late 90's.

Managing this growth demands tools for user augmented perception, memory, cognition, and communication. This paper outlines experience with intelligent tools that support the automated discovery of distributed experts and communities of expertise, the automated detection and tracking of emerging topics from unstructured multimedia data, and capabilities to increase organizational awareness (e.g., awareness of team members and materials in virtual collaboration environments). We first, however, introduce a knowledge management maturity model that frames our overall efforts. This article then describes the next stage beyond search engines to find knowledge, namely questions answering systems, and then describe systems created to access and collaborate with experts using the tools Expert Finder, XperNet and the Collaborative Virtual Workspace (CVW). Question answering systems combine natural language query understanding, information retrieval, information extraction

and answer generation technologies to provide users answers to questions. Expert Finder is an expert skill finder that exploits the intellectual products created within an enterprise to support automated expertise classification. XperNet addresses the problem of detecting extant or emerging areas of human expertise without a priori knowledge of their existence. Both Expert Finder and XperNet combine to detect and track experts and expert communities within a complex work environment. CVW (cvw.mitre.org) is a place-based collaboration environment that enables team members to find one another and work together. This article concludes with an outline of future research directions, notably in the area of automated question answering.

2 Knowledge Management Capability Maturity Model (KM-CMM)

The investigations described in this article are being explored in the context of a maturity model of knowledge management (KM) modeled after the Software Capability Maturity Model® (SW-CMM®) (www.sei.cmu.edu/cmm/). The Knowledge Management Capability Maturity Model (KM-CMM), summarized in Figure 1, describes the principles and practices underlying KM process maturity and is intended to help knowledge organizations improve the maturity of their knowledge processes in terms of an evolutionary path from ad hoc, chaotic processes to mature, disciplined KM processes. Like the SW-CMM, the KM-CMM is organized into five maturity levels:

1) *Initial*. At this level the knowledge process is characterized as ad hoc and occasionally even chaotic. Few processes are defined. Only a partial if any technical infrastructure to support knowledge discovery and sharing exists. As with the SW-CMM, success depends on individual effort and heroics.

2) *Repeatable*. At this level a basic knowledge management program is established to track requirements, content and investments. A necessary process discipline is in place to assure quality of knowledge and/or to repeat earlier knowledge successes and/or knowledge transfer on similar projects ensuring some basic knowledge quality and reuse. A knowledge focal point (KFP) is identified who is responsible for championing knowledge management efforts within the organization.

3) *Defined*. At this level the process for both knowledge management and knowledge engineering activities is documented. Processes are standardized and integrated across the organization. All projects use an approved, tailored version of the organization's standard process for developing and maintaining knowledge. Manual or semi-automated methods for mapping the organizations knowledge are applied and available across the enterprise. There exists knowledge management training and intergroup coordination of knowledge discovery and dissemination (e.g., via processes such as knowledge reviews and/or knowledge sharing exchanges).

4) *Managed.* At this level, detailed measures of the knowledge process and product quality are collected. Both the knowledge process and products are quantitatively understood and controlled.

5) *Optimizing*. At this level, continuous process improvement is enabled by quantitative feedback from the process and from piloting innovative methods, ideas, and technologies. Knowledge management activities are closely aligned with business functions.

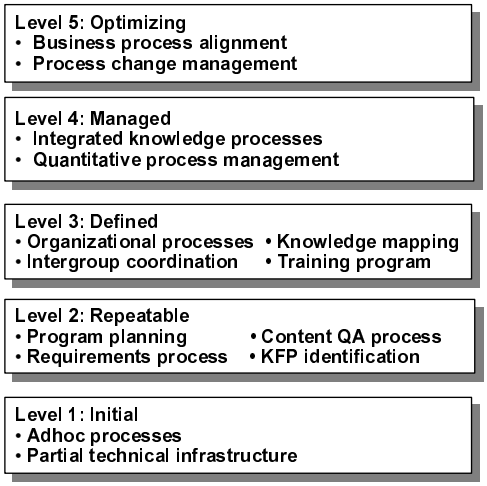

Figure 1: Knowledge Management Capability Maturity Model (KM CMM)

3 Question Answering

To enhance our own knowledge process maturity, one of the technologies we have been investigating is the use of automatic question answering. Question answering (QA) systems are an active current research area, including a TREC track on QA [Voorhees and Tice 2000] and a large US government program funded by ARDA on Advanced QUestion Answering for INTelligence [AQUAINT]. Question answering systems typically contain a few fundamental subsystems: question analysis, answer retrieval (which might include document retrieval and passage or fact extraction) and answer presentation generation. They often incorporate combinations of technologies such as information retrieval, information extraction, and language generation. Researchers are beginning to explore interactive QA, where users might have an opportunity to refine their questions or issue follow up questions. For example, at MITRE when our QA system called Qanda (Question AND Answering) [Breck et al. 2000] is given the question "Who was the architect of the Hancock building in Boston?" posed against a collection of five years of the LA Times, it retrieves the

statement "I.M. Pei was a student at M.I.T … He designed the John Hancock building in Boston."

Kinds of QA	Question/ Answer Complexity	Source Volume, Quality	Corpus, Resource Model	Answer Integration & Generation	Type of User Query	Type of System Answers
TREC Q&A On-line dictionaries, encyclopedias On line manuals (e.g,. UC) Web QA (e.g., Kiwilogic linguabot, e-Gain mail, Firepond) Multimedia QA	Moderate Q., Easy A.	small (100s MB), static, high quality source	Encyclo-pedic technical manuals	easy	FORM: - Keyword(s) - Phrase(s) - Question(s) TYPE: - Who - What - When - Where - How - Why - What-if INTENT: - request - command - inform	- Named entities - Phrase - Factoid - Link to document - Summary
	Easy to Moderate Q., Moderate A.	small to high (10 GB), dynamic, variable quality sources	Web	moderate		
	Hard Q., Hard A. [Interacts w/ Multilinguality]	very high, real-time, streaming, dynamic variable quality sources	Varied Multi-linguality (in Response and in Query)	hard		

Figure 2: Question Answering Characteristics

Figure 2 illustrates a range of question answering (QA) characteristics. For example, we can have QA from a selected document collection as in the Text Retrieval (TREC) QA track, retrieval of answers from semi-structured sources such as dictionaries, encyclopaedia or fact books, QA from massive, unstructured sources such as the web, and multimedia QA. As Figure 2 shows, there is a range of question/answer complexity, corpus volume, and degree of answer integration. Systems may address a variety of question forms (e.g., keyword, phrase, question) and types (e.g., who, what, why). Questions might encode a range of intentions such as a request for information, a command to perform some action such as a calculation, or also even information within the question (e.g., "What type of Titleist balls does Tiger Woods use?"). The answers might come in the form of a named entity, a phrase, a factoid, a link to a document or documents, or a generated summary. Additional characteristics include the degree of world knowledge in the system, its use of context and support for QA dialogue, if it has a user model and its nature (e.g., stereotypical, individualized, overlay), its task model, the structure of the domain, the degree of answer reuse in the system, and the degree of expected performance.

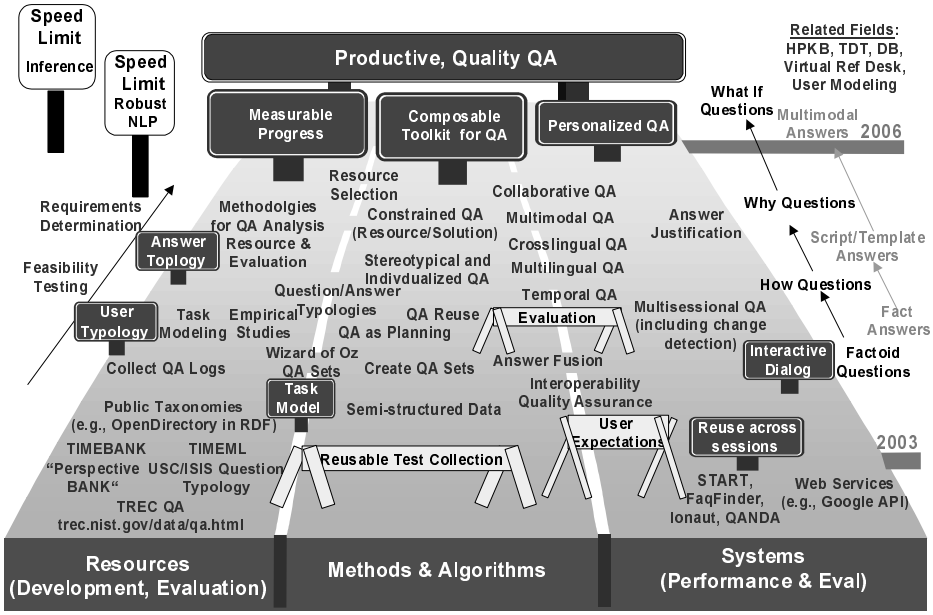

Figure 3: Question Answering Roadmap

Figure 3, illustrates a roadmap created at an LREC workshop in May of 2002 (www.lrec-conf.org/lrec2002/lrec/wksh/QuestionAnswering.html). This was produced to complement the existing ARDA QA roadmap available at www.nlpir.nist.gov/projects/duc/papers/qa.Roadmap-paper_v2.doc. The roadmap in Figure 3 is divided into three lanes dealing with resources necessary to develop or evaluate QA systems, methods and algorithms, and systems (including their performance and evaluation). The roadmap starts now and runs until 2006. Each lane leads to outcomes (indicated by sign posts) such as measurable progress from having shared resources, a composable QA toolkit, and personalized QA. An overall, long term outcome of QA systems that become high quality and enhance productivity. Sign posts along the road indicate intermediate outcomes, such as a typology of users, a topology of answers, a model of QA tasks (from both a system and user perspective), QA reuse across sessions, and interactive dialogue. Roadblocks along the way include the need to manage and possibly retrain user expectations, the need for reusable test collections and the need for evaluation methods. Overall workshop participants felt that general natural language processing and inference were limiters to progress, and so these were represented as speed limits signs on the left hand side of the road map. Here also we can see an arrow that indicates that feasibility testing and requirements determination are continuous processes along the road to productive, quality QA. On the right hand side of the road map we can see the progression of question and answer types. Questions progress from simple factoid questions to how to why then to what-if questions, whereas answers start out as simple facts but move to scripted or templated answers and then progress further to include automatically generated multimodal answers. Related fields such as high performance knowledge bases (HPKB), topic

detection and tracking (TDT), databases, virtual reference desks, and user modeling were noted as having particular importance for solving the general QA problem which will require cross community fertilization. Individual activities within the lanes are either currently planned or future desired events progressing toward longer term objectives.

4 Toward Multimodal Question Answering

A long range vision of ours is to create software that will support natural, multimodal information access, moving beyond written QA. As implied by Figure 4, this suggests transforming the conventional information retrieval strategy of keyword-based document/web page retrieval into one in which multimodal questions spawn multimodal information discovery, multimodal extraction, and personalized multimodal presentation planning. In Figure 4 the user of the future is able to naturally employ a combination of spoken language, gesture, and perhaps even drawing, eye movements, or facial expressions to articulate their information need which is satisfied using an appropriate coordinated integration of media and modalities, extracted from source media.

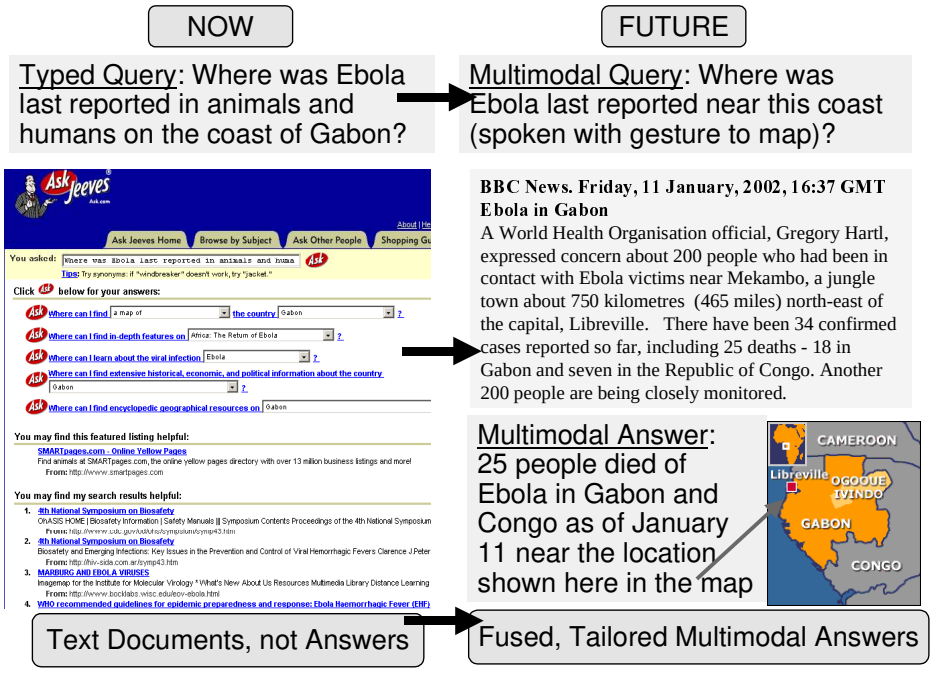

Figure 4: Ask Multimodal Questions, Get Multimodal Answers

As a step toward multimodal question answering, we have been exploring tools to help individuals access vast quantities of non-text multimedia (e.g., imagery, audio, video) [Maybury 1997]. Applications that promises on-demand access to multimedia information such as radio and broadcast news on a broad range of computing

platforms (e.g. kiosk, mobile phone, PDA) offer new engineering challenges. Synergistic processing of speech, language and image/gesture promise both enhanced interaction at the interface and enhanced understanding of artifacts such as web, radio, and television sources [Maybury 2000]. Coupled with user and discourse modeling, new services such as delivery of intelligent instruction and individually tailored personalcasts become possible.

Figure 5 illustrates one such system, the Broadcast News Navigator (BNN) [Merlino et al. 1997]. The web-based BNN gives the user the ability to browse, query (using free text or named entities), and view stories or their multimedia summaries. For example, the screen shot on the left of Figure 5 displays all stories about the Cuba from multiple North American broadcasts in June 2001. This format is called a Story Skim. For each story, the user can select a particular story and view story details (as in the left hand screen shot in Figure 5), including a closed caption text transcription, extracted named entities (i.e., people, places, organizations, time, and money), a generated multimedia summary, or the full original video.

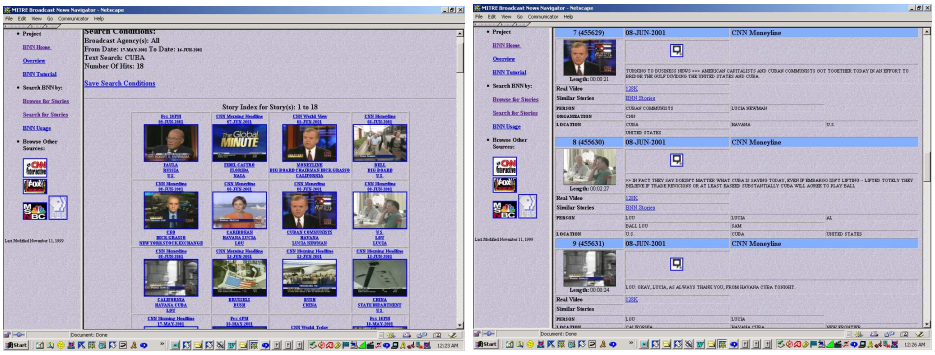

Figure 5: Tailored Multimedia News Story Skim (Left) and Story Detail (Right)

In empirical studies, [Merlino and Maybury 1999] demonstrated that users enhanced their retrieval performance (a weighted combination of precision and recall) when utilizing BNN's Story Skim and Story Details presentations instead of mono-media presentations (e.g., text, key frames). Figure 6 illustrates how users could obtain performance close to that of dealing with the original video source, except the multimedia presentations of key frames and named entities could be searched about two to three times as fast as manual search through linear video. In addition to performance enhancement, users reported increased satisfaction (8.2 on a scale of 1 (dislike) to 10 (like)) for mixed media display (e.g., story skim, story details).

Just as users were show to be more effective when given mixed media presentations, we also found higher computer algorithm performance on media analysis and segmentation using multimedia cues from audio, video, and close caption sources to determine commercial start/start, classification of shots (e.g., anchor, reporter, commercial), and transitions from one state to another (e.g., anchor to reporter in the field). We utilized simple annotation tools allowing non-experts to markup a corpus of video for features such as program start/stop as well as

commercial and story segments. Then we automatically induced a cross-modal statistical model for video segmentation and transition detection using hidden Markov models. Current efforts are investigating issues such as automatically discovering users topical interests and media preferences by monitoring their queries and interaction with the system in order to dynamically either search for information for them or tailor retrieved information to their preferences.

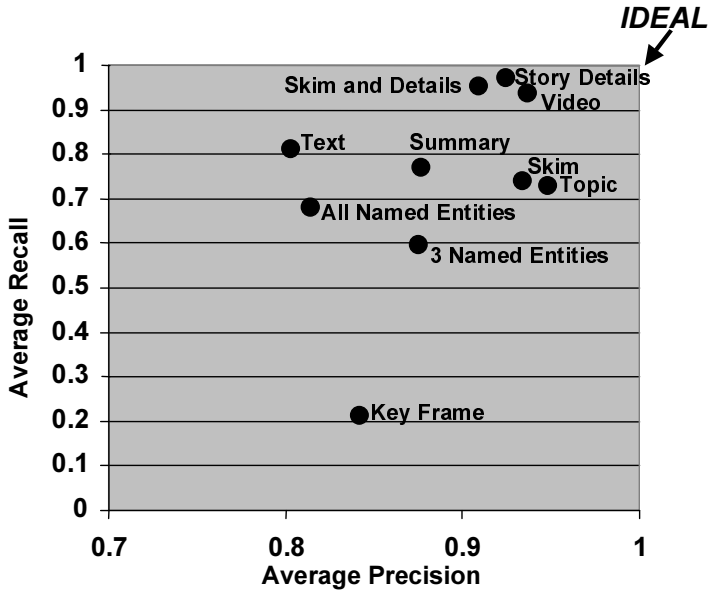

Figure 6: Relevancy Judgement Performance with Different Multimedia Displays

5 Expert Finding and Expert Network Discovery

Just as information on demand is important, so too its essential to be able to find expertise on demand. Unfortunately, resumes and manually populated skills databases are well known to be uneven, out of date, or simply non existent in many organizations. We have created a system that analyzes user created documents and mentions of experts in newsletters to automatically construct a keyword profile of a user's expertise. Expert Finder [Mattox et al. 1998, 1999; Maybury et al. 2001] as the system is called, looks at products produced by expert (e.g., briefings, papers, web pages) and products that mention the individual (e.g., Newsletters, articles in magazines). In the latter case, information extraction software (using NameTag from IsoQuest Corporation) is used to extract the individual's name which is then correlated with topics mentioned in the surrounding text. The more documents linking the individual to a topic and/or mentions of the individual with a topic, the higher the expertise rating of the individual. Additional weight is given a resume. Figure 7 illustrates Expert Finder in action in which a query for "machine translation" has return a rank ordered list of experts within the company, drawing upon evidence from

employee publications, mentions in corporate communications, and project leadership information. In an empirical evaluation, when searching for the top five experts in an area, Expert Finder was able to automatically retrieve over 30% of the experts human experts would recommend manually.

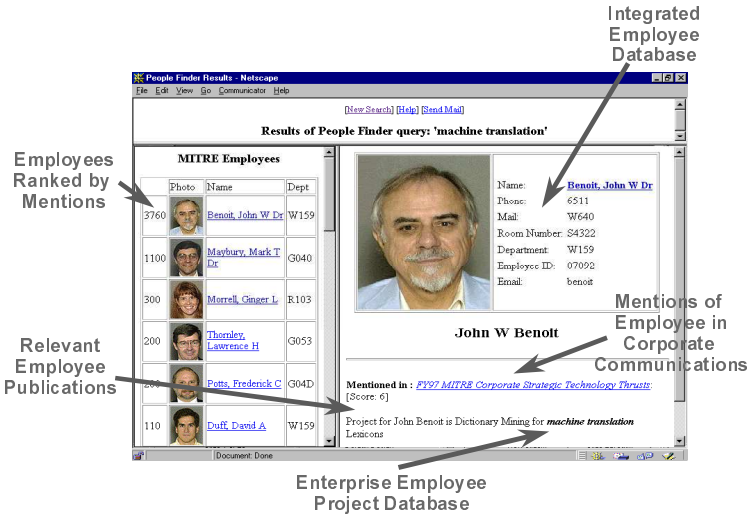

Figure 7: Expert Finder

In contrast, in separate research we seek to identify networks of experts. XpertNet works without user queries to identify expertise areas; a distinction between it and other expertise locator tools. 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. We expect domain-specific expertise models in niche technology areas (e.g., Perl programming). An example of an expertise network, with individual identities masked out for privacy, is provided in Figure 8. In this "map", nodes represent people within the organization that are "involved" in our natural language processing work. We use shape to relate to the technical skill rating (organizationally assigned) of each network member (e.g., double box refers to personnel with a level 5 rating). We can also use other designators such as labels or colors to indicated individuals in the same organization. We are presently engaged in an planning for an enterprise wide roll out of an expertise management solution.

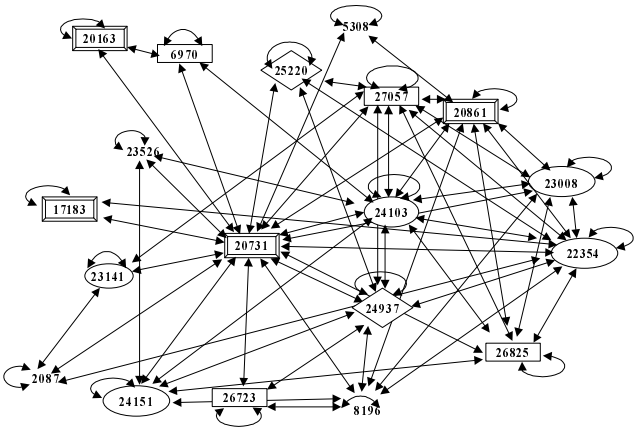

Figure 8: An Expertise Network

6 Human-Human Collaboration

Just as it is important to provide mechanisms for multimodal human machine collaboration, so too it is important to enable multimodal human human collaboration, augmenting current face-to-face interactions. Figure 9 graphically depicts the importance of team efforts and attempts to relate several levels of human collaboration, which build upon one another. Levels range from awareness of individuals, groups and activities, to sharing information with one another, to coordinating individual activities, to working jointly together, ultimately leading up to shared intent.

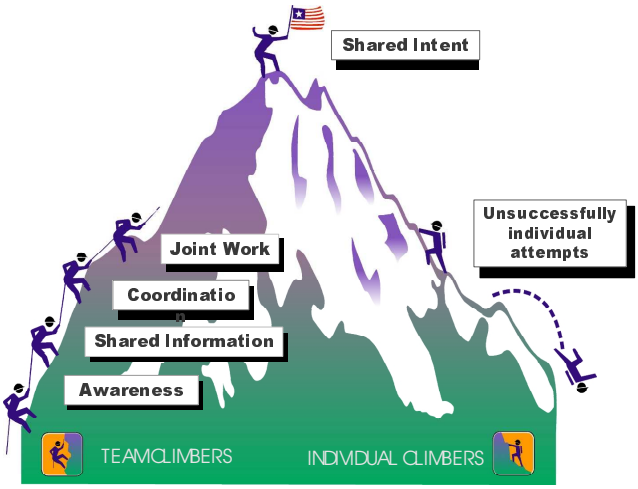

Figure 9: Levels of Collaboration

As detailed in Table 1, each of these levels of interaction implies different activities, classes of tools and associated media and modalities. For example, basic awareness of others, their communication capabilities (e.g., text, audio, video), availability, and perhaps even their activities is a fundamental prerequisite to collaboration. Tools such as electronic calenders, publish/subscribe mechanisms, presence information, and expertise finding tools can facilitate this awareness. Communication of awareness information typically occurs using text, graphics, and audio or visual alerts.

At the next level users can share information with one another at conferences, workshops, tutorials or just using personal communication in electronic mail, chat or video teleconference. Users can go beyond information sharing to coordination, the next level, which might involve creating shared assessments or shared plans in group brainstorming or decision meetings, possibly supported by decision support tools. Coordination might rely upon many media and modalities.

Joint work can occur face-to-face but can also be mediated by tools such as shared whiteboards or shared applications which can capture user preferences and application interactions. Workflow tools can facilitate sequencing and controlling interdependent efforts. Finally, building upon all of the underlying levels, the establishment of shared intent in a relationship typically grows over many, often face-to-face, interactions.

LEVEL	Activities	Tools	Media
Shared Intent	• Shared Purpose • Co-dependent	• Strategic Alliances	• Face-to-face
Joint Work	• Shared goals • Joint goal creation • Cross-organizational teams	• Workflow • Whiteboard • Shared applications	• Application actions • Gesture • Text • Audio • Video
Co-ordination	• Shared plans • Group meetings	• Decision Support • Brainstorming tools	• Text • Audio • Video
Shared Information	• Meetings • Conferences • Briefings and presentations • Training	• E-mail, chat, VTC • Web pages, Portals • Publications	• Text • Messages • Audio • Video
Awareness	• Shared calendars • Shared presence	• Electronic calendars • Publish/subscribe • Alerts • Presence • Expert finding	• Text • Graphics • Audio • Video

Table 1: Collaboration Levels, Example Tools, and Media

For a number of years we have been exploring human human group collaborations within distributed, virtual environments. Our work has resulted in the open source software (cvw.sourceforge.net), Collaborative Virtual Workplace (CVW), a screenshot of which is shown in Figure 10. CVW incorporates a comprehensive suite of tools that support many of the tasks outlined in Table 1, including shared whiteboarding, audio/video/text conferencing, user presence awareness, access

control, and persistent virtual spaces (i.e., virtual rooms which contain applications, documents, and users).

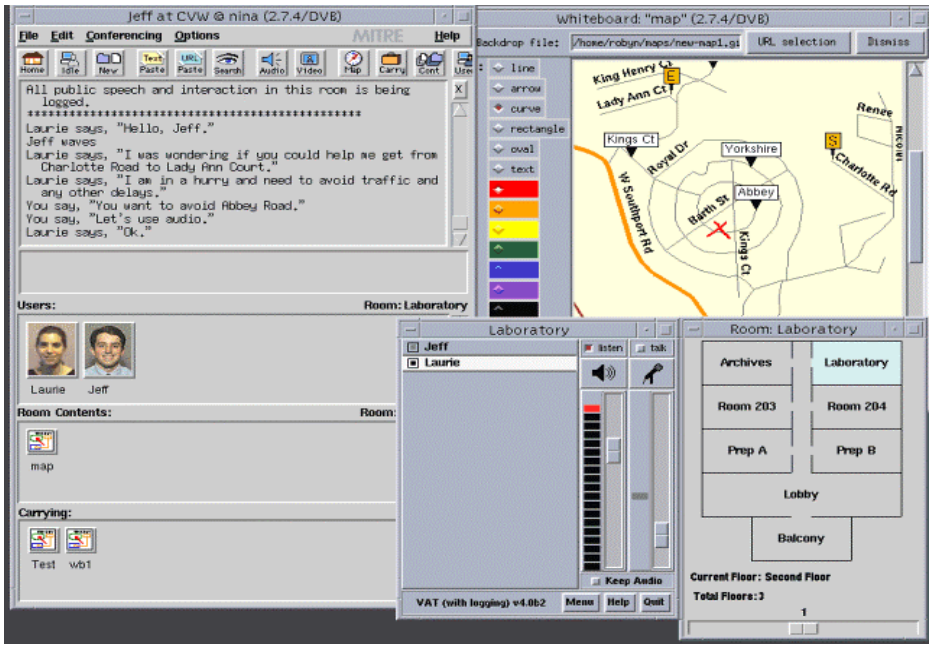

Figure 10: Collaborative Virtual Workplace

[Maybury 2001] describes the functionality and operational use of this place-based environment by hundreds and thousands of users in two major organizational settings for analysis and planning. In order to understanding the operational impact and evaluate the effectiveness of these tools, as well as to understand technical infrastructure issues, we have found it essential to instrument user activities within these virtual environments. We have used MITRE's multimodal logger to accomplish this, which we describe next. [Hall 2000] details methods for measuring impacts in several collaboration technologies within several organizational.

7 Multimodal Logging and Evaluation

MITRE's Multimodal logger [Bayer et al. 1999] supports the recording, retrieval, annotation and visualization of data collected in human-computer and human-human interactions. The Multimodal logger incorporates a database structure which groups datapoints by application (e.g., audio utterance, text chat, whiteboard use, video conference) and applications by session. It supports the typing of data points via MIME types, provides an easy-to-use API for instrumenting existing applications and tools for reviewing and annotating data collected via instrumentation.

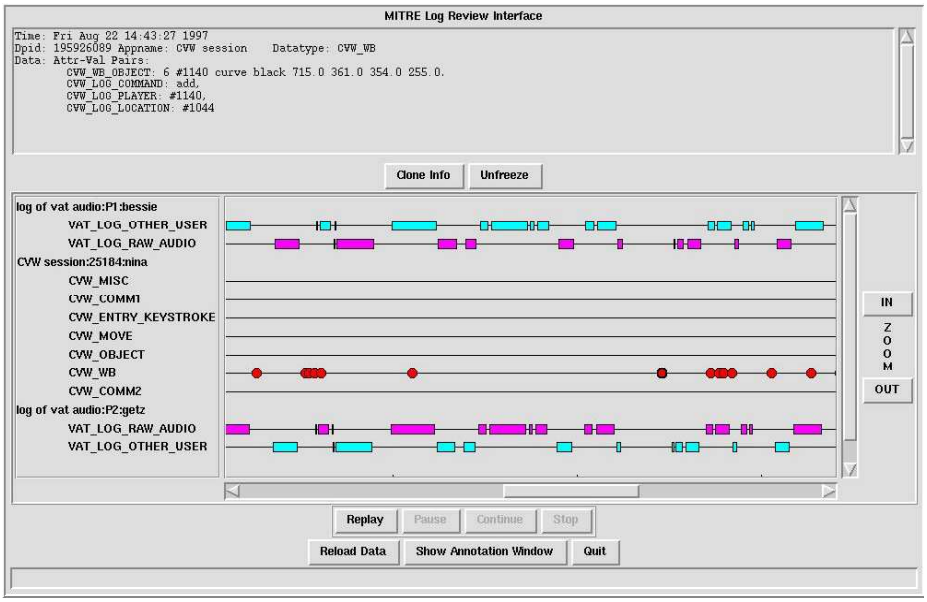

Figure 11:. Multimodal Logging and Annotation

Figure 11 illustrates the visualization of multimedia events across a range of applications such as whiteboarding (CVW_WB), start, end and duration of events in audio conferencing (VAT), movements among virtual rooms (CVW_MOVE) and object manipulation (CVW_OBJECT). The user can zoom in or out to inspect specific events as well as add further annotations to this automatically constructed event log. This supports analyses, for example, of multiparty communication to look at properties such as frequency of user communications and actions, discourse events such as interruptions, and cross modal events such as co-occurring speech and gestures.

DARPA's Intelligent Collaboration and Visualization initiative (zing.ncsl.nist.gov/nist-icv) utilized MITRE's multimodal logger in support of collaboration system evaluation. Working initially with NIST, NIMA and CMU, MITRE developed an assessment methodology for collaboration systems [Cugini et al. 1997, Damianos et al. 2000] that includes a framework of four levels of abstraction. A requirements level captures the work and transition tasks to be performed, and the social protocols and characteristics of the group performing the tasks; the next level specifies the capabilities (e.g., shared workspace, communications, etc.) required to perform the work; a services level describes specific services (e.g., text chat, whiteboard) that could be used to deliver the capabilities, and a technology level describes specific implementations of services. Associated with each level are appropriate assessment metrics. Assessments can be made at multiple levels of this framework, depending on the intended needs of the evaluators, whether they are users, researchers, or systems designers. Community defined multimodal evaluations are essential for progress, and that the key to such progress is a shared infrastructure of

benchmark tasks, evaluation tools, and training and test sets to support cross-site performance comparisons.

8 Conclusion

In conclusion, we have shown how intelligent information access tools such as question answering and news understanding can enhance human cognitive performance. Moreover, we have illustrated new tools to detect expertise automatically from intellectual products and to discover networks of experts. Virtual place-based collaborative environments can both support these communities and can be exploited to invoke groups of experts to perform joint tasks. Finally, we have described a capability maturity model for knowledge management that provides a framework for organizations to measure and manage their levels of capability in this strategic area.

Aknowledgement

I give special thanks to the LREC question answering participants for their roadmap contributions. I am indebted to David Mattox for his assistance with the creation of MITRE's Expert Finder. We would also like to thank Inderjeet Mani, whose research efforts supported the development of this application, and Chris Elsaesser for the original idea for an expert finder. I thank Ray D'Amore for his social network analysis vision and Manu Konchandy for his contribution to the development of XperNet tools. Marc Light and John Burger are responsible for QANDA. I thank the former CVW prototyping team. I'm indebted to Tamra Hall for her initial inspiration regarding levels of collaboration. I thank Jean Tatalias and her corporate knowledge management team for their continued ideas about and insights into knowledge management.

References

[AQUAINT]. Advanced Question Answering for Intelligence (AQUAINT) www.ic-arda.org/InfoExploit/aquaint.

[Bayer et al. 1999] Bayer, S., Damianos, L., Kozierok, R., Mokwa, J. 1999. The MITRE Multi-Modal Logger: Its Use in Evaluation of Collaborative Systems. ACM Computing Surveys. March. www.mitre.org/technology/logger.

[Breck et al. 2000] Breck, E., Burger, J. D., Ferro, L., House, D., Light, M., Mani, I. 2000. A Sys called Qanda. In Vorhees, E.. and Harman, D. The Eighth Text Retrieval Conference, NIST Special Pubs. February 2000. 499-506.

[Cugini et al. 1997] Cugini, J., Damianos, L., Hirschman, L., Kozierok, R., Kurtz, J., Laskowski, S., and Scholtz, J. 1997. Methodology for Evaluation of Collaboration Systems. zing.ncsl.nist.gov/nist-icv/documents/method.pdf.

[Damianos et al. 2000] Damianos, L, Drury, J., Fanderclai, T., Hirschman, L., Kurtz, J. Oshika, B. 2000. Evaluating Multi-party Multi-modal Systems. In Proceedings of LREC 2000, vol. III, 1361-1368.

[Hall 2000] Hall, T. July 2000. A Practioners Guide to Evaluating Collaboration Systems. Community Collaboration Facilitation Center (CFC). collaboration.mitre.org/practguide/PractionersGuide.html

[Mattox et al. 1998] Mattox, D., Smith, K., and Seligman, L. 1998. Software Agents for Data Management. In Thuraisingham, B. *Handbook of Data Management*, CRC Press: New York. 703-722.

[Mattox et al. 1999] Mattox, D., Maybury, M. and Morey, D. 1999. Enterprise Expert and Knowledge Discovery. International Conference on Human Computer International (HCI 99). 23-27 August 1999. Munich, Germany. 303-307.

[Maybury 1997] Maybury, M. T. (ed.) 1997. *Intelligent Multimedia Information Retrieval*. Menlo Park: AAAI/MIT Press. www.aaai.org/Press/Books/Maybury2/maybury2.html

[Maybury 2000] Maybury, M. Feb. 2000. News on demand: Introduction. *Communications of the ACM*. 43(2): 32-34. www.acm.org/cacm/0200/0200toc.html

[Maybury 2001] Maybury, M. December 2001. Collaborative Virtual Environments for Analysis and Decision Support. *Communications of the ACM* 14(12): 51-54. In Ragusa, J. and Bochenek, G. (eds). Special Section on Collaboration Virtual Design Environments. www.acm.org/cacm/1201/1201toc.html

[Maybury et al. 2001] Maybury, M., D'Amore, R, and House, D. December 2001. Expert Finding for Collaborative Virtual Environments. *Communications of the ACM* 14(12): 55-56. In Ragusa, J. and Bochenek, G. (eds). Special Section on Collaboration Virtual Design Environments. www.acm.org/cacm/1201/1201toc.html

[Merlino et al. 1997] Merlino, A., Maybury, M. and Morey, D. 1997. Broadcast News Navigation Using Story Segments, ACM International Multimedia Conference, Seattle, WA, November 8-14, 381-391.

[Merlino and Maybury 1999] Merlino, A. and Maybury, M. 1999. An Empirical Study of the Optimal Presentation of Multimedia Summaries of Broadcast News. Mani, I. and Maybury, M. (eds.) *Automated Text Summarization*, MIT Press, 391-402.

[Morey et al. 2000] Morey, D.; Maybury, M. and Thuraisingham, B. editors, Fall 2000. *Advances in Knowledge Management: Classic and Contemporary Works*. Cambridge: MIT Press.

[Voorhees and Tice 2000] Voorhees, E. and Tice, D.M. 2000. The TREC-8 Question Answering Track Evaluation. Voorhees, E. and Harman, D.K. (eds.) *The Eighth Text Retrieval Conference (TREC-8)*, NIST Special Publication 500-246, 83-106.