

# Knowledge Management for Research Communities

Brian R. Gaines and Mildred L G Shaw

University of Calgary

Calgary, Alberta, Canada T2N 1N4

{gaines, mildred}@cpsc.ucalgary.ca, <http://ksi.cpsc.ucalgary.ca/KSI/>

From: AAAI Technical Report SS-97-01. Compilation copyright © 1997, AAAI ([www.aaai.org](http://www.aaai.org)). All rights reserved.

## Abstract

Increasing ease of access to the Internet is making it feasible for geographically dispersed communities to work closely together, coordinating their activities through electronic mail, digital document archives, and access to remote computing facilities. This article reports on practical experience of knowledge management for a number of communities and projects. It gives an overview of the technologies available, their accessibility, ease of use and impact on collaborative activities. It focuses on the practical problems that arise, the limitations of existing technologies, and how these may be overcome. Examples are given of projects encompassing a range of multimedia digital technologies from list servers, through World Wide Web document archives, production in a few days of CD-ROMs containing movies and digitized foils and documents giving a complete account of working meetings, to active documents that are themselves expert systems.

## 1 Introduction

The Knowledge Science Institute (KSI) was set up in 1985 with a dual mandate: to study and model the knowledge economy (Bell, 1973) with a view to forecasting trends and modeling knowledge processes; and to innovate in selected areas that might lead to insights into the impact of knowledge support systems (Gaines, 1990). Figure 1 shows the major research programs addressing these objectives:

- *Information systems infrastructure* modeled as a tiered structure of learning curves for electronic devices, computer architecture, software engineering, interactive systems, knowledge representation and inference, knowledge acquisition, autonomous systems, and socially organized systems (Gaines, 1991b).
- *Knowledge dynamics* modeled in social, psychological and logical terms as a process of distinction creation, dissemination, application and evaluation (Gaines, 1989; Gaines and Shaw, 1993a).
- *Collective stance model* of the formation of expertise through positive feedback processes in the human species viewed as a single organism recursively partitioned in space and time into sub-organisms that are similar to the whole; these parts include societies, organizations, groups, individuals, roles, and neurological functions (Gaines, 1994b).

- *Knowledge acquisition and representation* methodologies and tools for eliciting, modeling and reasoning with human knowledge (Gaines and Shaw, 1993a,b).
- *Multimedia active documents* supporting the dissemination of knowledge through systems integrating text, diagrams, pictures, movies, sound, concept maps, semantic nets, animation, simulation and access to general information systems (Gaines and Shaw, 1993c).
- *Groupware knowledge support systems* targeted on the systematic acceleration of scientific research through networked access to knowledge-based systems and multimedia active documents (Gaines and Shaw, 1993c, 1995a; Gaines, Norrie and Lapsley, 1995).

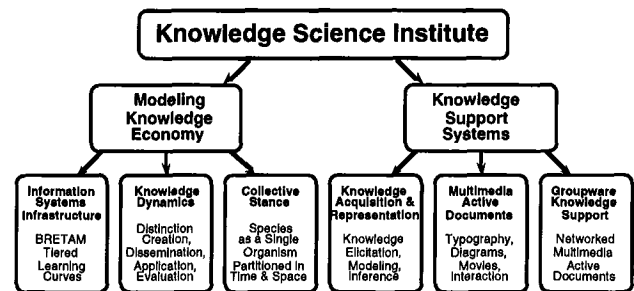


Figure 1 Knowledge Science Institute research program

One conclusion of the KSI studies is that a major anachronism in human knowledge processes has, until recently, been the continued reliance on paper media for knowledge dissemination (Gaines, 1993). While book production technology has gone through several major technological advances, the end-product is basically the same as it was for Gutenberg and Caxton over five hundred years ago. Similarly, scholarly journals have changed little in their presentation and operation since the inception of the *Philosophical Transactions* of the Royal Society in March 1665. When one considers the limitations of the paper medium in communicating knowledge, and contrasts it with the capabilities of current multimedia computer and communication technology, the incentives to make major changes in scholarly publishing are very high.

Books and journals primarily support typographic text and monochrome diagrams and photographs, and require separate indexing for purposes of contextual search. Computer-readable multimedia documents support the

same quality of typographic text, color diagrams and photographs, sound and video, animation, simulation, computation, and are intrinsically contextually searchable. They can be produced and distributed on CD-ROM at a substantially lower costs than paper books or journals, and can be read on personal computers and selectively printed on laser printers. They can also be produced and distributed on the Internet, providing an active medium supporting many forms of scholarly discourse.

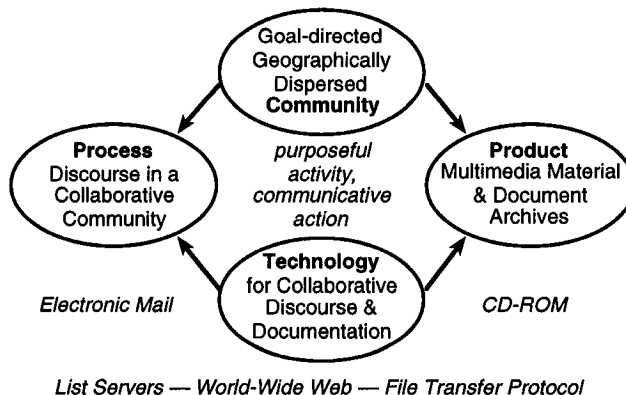
The objective of the KSI research program reported in this article is to provide a next generation of open architecture knowledge support systems (Gaines, 1994a) supporting collaboration in scholarly communities through information technology with a view to achieving systematic acceleration of human knowledge processes. Specifically, a new layer of knowledge level tools has been developed that complements existing list servers and electronic document archives with research coordination facilities. The tools are designed to support the knowledge processes of scientific communities through active document, hypermedia, concept mapping, knowledge representation and active agent technologies. They use existing Internet protocols for mail, file transfer and hypermedia to integrate with related systems. They provide import and export facilities for common document formats to integrate with existing document production. The data formats are designed for parallel use on CD-ROM to provide alternative distribution.

The pragmatic objective is to achieve widespread use of greatly improved facilities for scientific discourse and collaborative project management. A deeper objective is to study the impact of the proposed technologies on actual research practice with a view to designing a next generation of tools targeted on the changes in patterns of research behavior already apparent as new information technologies have come to play a significant role in scholarship (Manitoba, 1993; Renear and Bilder, 1993).

Figure 2 shows the general situation and support technologies studied in this article. A goal-directed, geographically dispersed community is engaged in a purposeful activity involving *communicative action* (Habermas, 1981)—that is, discourse in which mutual understanding is a joint objective. The *process* of discourse in a collaborative community can be supported through technologies such as electronic mail (email) and mailing list servers. The *product* of this discourse is captured in multimedia materials and document archives whose dissemination can be supported through technologies such as CD-ROMs and the Internet file transfer protocol (FTP). Hybrid client-server technologies such as those of the World Wide Web may be used to support both process and product through structured discourse with archives indexed and linked to a record of the discourse.

The principles and applications of the knowledge-level tools developed by the KSI have been reported in a

number of articles which are available in publications (Gaines and Shaw, 1993c, 1994a,b; Kremer and Gaines, 1994; Shaw and Gaines, 1994) and through the web (<http://ksi.cpsc.ucalgary.ca/articles/>). This article illustrates a variety of techniques of knowledge management for research communities.



**Figure 2 Supporting the knowledge processes of a dispersed community**

## 2 Background to Studies—IMS GNOSIS Test Case

To provide a context for the projects used as examples in this article, it is relevant to give a brief overview of the international GNOSIS community whose support has been a focus of the major projects. IMS TC7 'GNOSIS', was one of 6 one-year test cases under the international Intelligent Manufacturing Systems research program which started in the second quarter of 1993. The project (GNOSIS, 1994) involves over 100 participants in 31 industry and university organizations in 14 countries, with the objective of developing a *post mass production manufacturing paradigm* involving *reconfigurable artifacts*. The project has made extensive use of electronic mail and electronic document archives to coordinate its activities, and the studies reported are part of an investigation to improve such coordination in the main 10-year study commencing in 1995.

Figure 3 shows the partners involved in the GNOSIS consortium together with their geographic locations. Figure 4 shows the meetings during the first year of the project and the cost impact of coordinating such a dispersed project—that meetings accounted for some \$1.1M of the estimated \$5M budget for the 1-year test case. The meetings and associated costs were probably necessary in the initial stages of a major collaborative venture involving partners many of whom had never previously interacted. However, the costs involved draw attention to two considerations relevant to multi-media communication:

- Each major meeting should be treated as an expensive resource and knowledge management techniques should be used to maximize the value received. In particular, the discourse at meetings should be recorded, indexed and archived, as should the slides presented, documents circulated, and so on.
- Other knowledge management media such as telephone, tele- and video-conferencing, fax, email, list servers, document archives, and so on, should be used effectively to reduce the need for physical meetings.

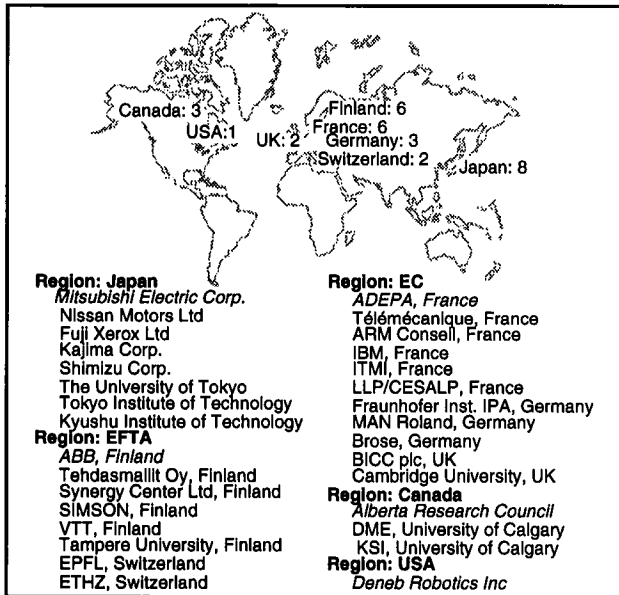


Figure 3 The GNOSIS consortium

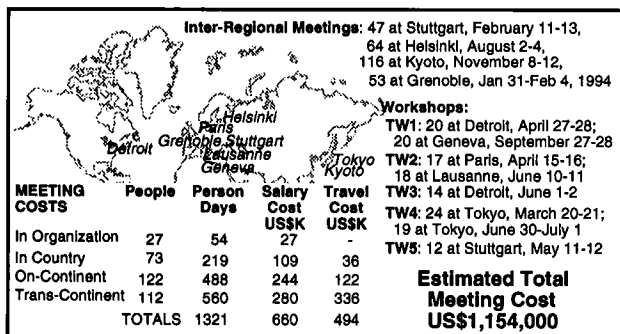


Figure 4 GNOSIS meeting costs in first year of operation

### 3 Digital Movies as Tutorial Demonstrations

One of the common activities at the initial GNOSIS meetings was for partners to give demonstrations to one another of ongoing research in their laboratories. The

typical form of demonstration is a presentation of some system in operation together with a commentary on what is happening, why, and how this illustrates the overall objectives and state of the research. This type of tutorial demonstration is ideally captured as a movie which can be edited, digitized and then disseminated over the Internet or through CD-ROM.

Digital movies are typically expected to be large data structures with poorer quality reproduction than analog tape or video-disk. However, for computer demonstrations this is not the case, and movies may be made that are accurate in reproducing the details of the computer screen while occupying little disk space. This section details some practice in experience in producing such a movie with a resolution of 512 by 342 pixels, high-quality sound, playing for 8 minutes and occupying only 11MBytes. Figure 5 shows a frame from the movie being played in Apple's *MoviePlayer*. It shows interaction in a word processor with a document that itself contains an embedded QuickTime movie. The size and resolution of the demonstration are apparent in the quality of the text and images shown.

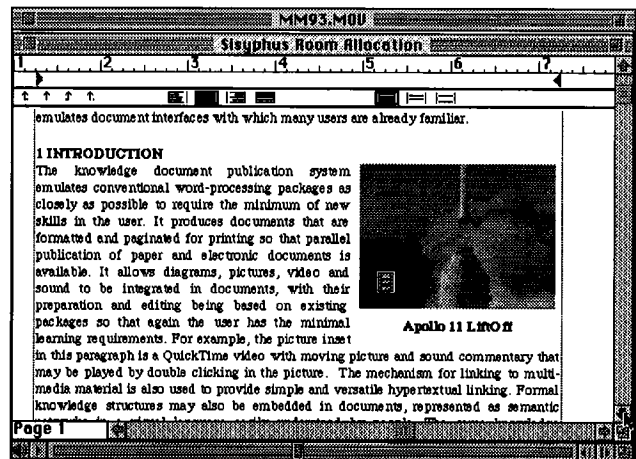


Figure 5 Tutorial demonstration as a high-resolution QuickTime movie

The movie is a tutorial demonstration of knowledge acquisition, active document and expert system technology developed in the KSI, all technologies that are relevant to GNOSIS and were demonstrated at various workshops. It was originally prepared for the CD-ROM publication of the ACM SIGGRAPH MultiMedia93 Conference, Anaheim, California, August 1993, and is a demonstration of technologies described in a paper in the proceedings (Gaines and Shaw, 1993c). The movie demonstrates that knowledge management for tutorial demonstrations can be achieved through the creation and dissemination of digital movies.

## 4 Capturing a Workshop in a Multimedia Proceedings

Formal workshops and conferences conventionally put much effort into producing a proceedings in a standard format specified to the authors so that they can produce camera-ready copy. However, modern optical character recognition (OCR) software coupled with style sheets in a word processor can be used to create a high-quality proceedings from original documents in arbitrary and varied formats. In addition, presentation slides can be digitized and presentations and discussion videotaped and digitized to create multimedia material that can be included in the electronic version of the proceedings.

As an example, a camcorder was used to capture software demonstrations and presentations at a 3-day GNOSIS workshop in Tokyo in March 1993 involving some 30 participants. The slides presented were collected together with the handouts circulated, largely technical papers from conference proceedings and internal reports. On return to Calgary, the papers were optically recognized using Xerox *AccuText* with an accuracy of some 98%, remarkably high considering that some of the papers were n'th generation copies of double-column conference material in 9pt type. The slides were digitized and the movies converted to QuickTime and lightly edited in Adobe *Premiere*. The movies were shot with a Sony *TR101* which has optical image stabilization that compensates for the slight instabilities of long hand-held shots. The editing was necessary primarily to allow for participants walking in front of the camera in what was a very informal setting—it is not appropriate to attempt to establish studio conditions at a scientific workshop.

The resultant material was put into a uniform style in Microsoft *Word*, resulting in a 300-page workshop proceedings (GNOSIS, 1993) containing 30 articles and 8 QuickTime movies. Figure 6 shows a page from the proceedings with the heading, comments, and movies of computer software demonstrations preceding a technical paper. The production time to OCR the source material, edit it into a uniform format, digitize the slides and movies, and issue the proceedings in print and CD-ROM was one week. The effort required was some 40 person-hours. Thus, a high-quality workshop proceedings can be produced *after* the event without any additional effort on the part of the participants, indeed without any advance warning that there is any intent to do so. The total cost of producing the proceedings was about \$3,000, about the same as the air fare to travel to Tokyo.

This application is significant to knowledge management for the semi-formal discourse in scholarly communities that takes place through seminars, workshops and informal conferences. Much of what occurs at such meetings is currently lost and not disseminated to others, and yet digital document technology now makes it simple to capture the essence of a meeting without imposing constraints on the participants.

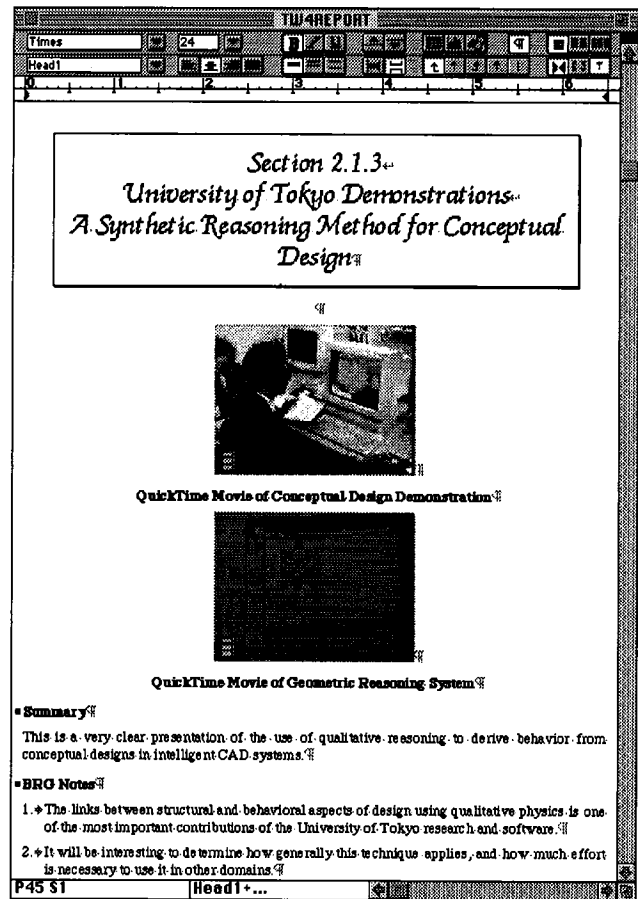


Figure 6 Workshop proceedings with embedded demonstrations

## 5 Concept Maps Indexing Research Archives

At the end of the first year of the GNOSIS project, the project reports, slide sequences, photographs, movies, and so on, were issued on a CD-ROM. To enable widespread access to the reports, a hybrid format was used that could be read on Macintosh, Windows and Unix platforms, and all the reports were issued in Microsoft *Word*, Farallon *Replica* and Adobe *PostScript* formats. The total volume of material on the CD-ROM was: 57 reports totaling 1590 pages, each in 3 formats; 11 movies totaling 70 minutes; plus software and maps of the material. 3 test masters were made of the CD-ROM using a Philips *CDD521* recorder and OMI *QuickTOPIX* software. The third one was sent to a CD pressing company that produced 500 copies in 10 days at a cost of \$1500. The second test master was very close to the final one and was sent off at the same time to a Japanese partner in GNOSIS who was scheduled to demonstrate the GNOSIS materials at a major IMS conference shortly after the production versions would be available.

The combination of high-speed production of one-off CD-ROMs and low-cost production of quantity CD-ROMs provides an extremely powerful and flexible technology for the knowledge management of technical material. In particular, the space required to archive CD-ROMs is very much less than that of the equivalent paper documents, and the cost of mailing them is also very much lower. Additionally, the CD-ROM can contain colored diagrams, photographs, sound movies, simulations and other software, that it would be very much more expensive to issue in other formats.

One factor significant to the usability of a CD-ROM is the type of indexing provided. It is also very relevant to the CD-ROM as a presentation medium. The 'corporate image' projected by access to material through a simple Unix or Windows directory is not the same as that projected by access through some multimedia navigational tool that provides a user-friendly contextual map of the material. There are commercial tools that index document collections by content and provide an information retrieval interface to the material. These are appropriate to reference materials that need to be searched by content. However, they again present a rather 'technical' and unattractive interface, and even when they are used it is appropriate to provide higher-level maps of the material.

For the GNOSIS archives, it was appropriate to use as an indexing tool *Mediator* (Gaines et al., 1995), a knowledge management system that had been developed to support collaborative activities across the network as part of the GNOSIS research program. The *Mediator* implementation was based on groupware concept-mapping tools that were already in use for indexing multi-media materials (Gaines and Shaw, 1995a). Figure 7 shows the GNOSIS project archives being accessed through layered concepts maps. The map in the window at the upper left is a top level "Server Agent" that manages a particular collection of material. In the example shown a local user is accessing material directly through this agent. Remote users connect to the server agent over the network using client agents that give them the same functionality through calls to the server.

The concept map at the top left is currently write-disabled, and the cursor has changed to a button as the user mouses over the "Group Photo" node. Clicking at this point will display the photograph in a separate window. The user has already clicked on the node "GNOSIS Final Reports" to open the concept map shown at the lower right. This has a node for each report, and clicking on one will open the appropriate report, in this application using Farallon's *Replica*. The node at the top left gives access to a series of slides on the project displayed using *Replica*. A similar node in the original concept map at the top left gives access to a movie on the GNOSIS project that will be opened in Apple's *MoviePlayer*.

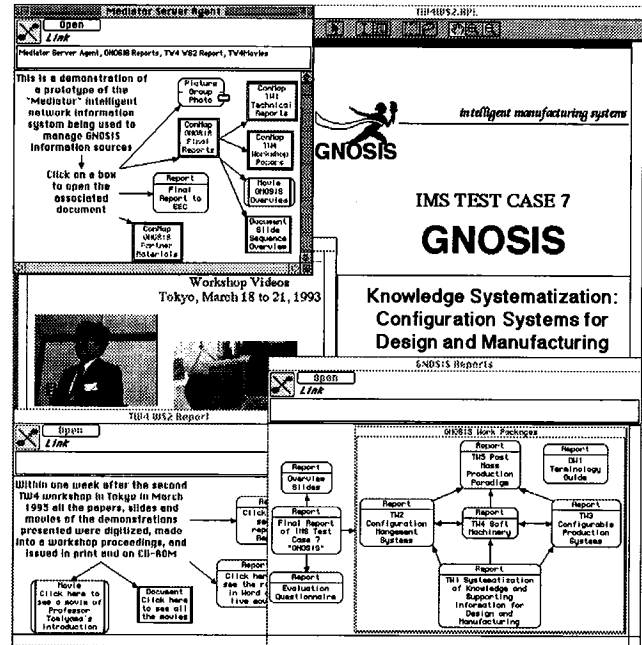


Figure 7 Accessing the GNOSIS archives through layered concept maps

The user has already clicked on the "TW4 Workshop Papers" node in the concept map at the top left, and opened the relevant concept map at the bottom left. She has then clicked on the node "Click here to see the report in Replica", and opened the report visible at the back on the right of Figure 7. She has also clicked on the "Click here to see all the movies" node and opened the KWrite document visible behind the concept maps. This displays eight QuickTime movies of various demonstrations given at the Workshop, any of which can be played by double clicking on it.

The concept mapping tools have been ported to operate through the World Wide Web so that the archives may be accessed through the Internet (Gaines and Shaw, 1995a,b). They have also been ported to Java so that they may be edited in Web browsers (Flores-Mendez, 1997) and provide a versatile knowledge management system on the World Wide Web.

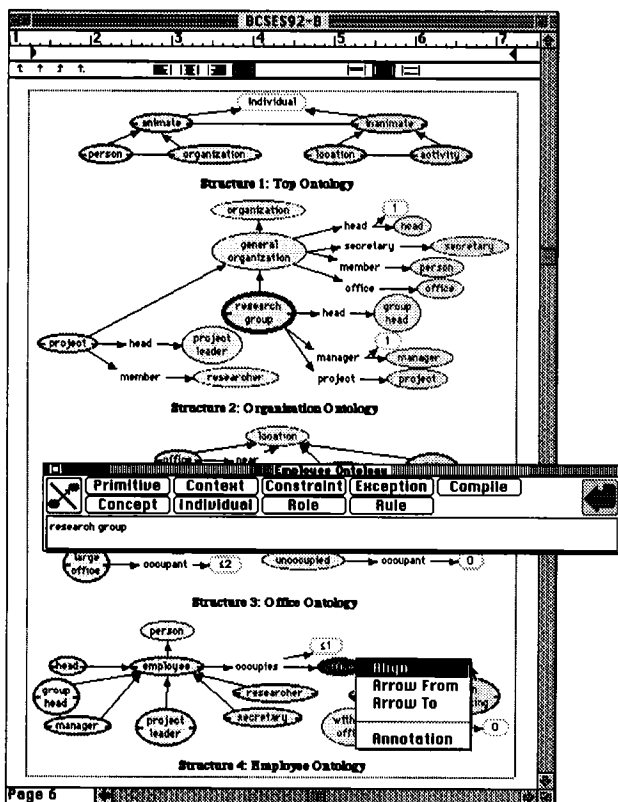
## 6 Active Documents

KWrite (Gaines and Shaw, 1993c) is a document processing system that supports embedded active objects including concept maps. The embedded panes may be used to display graphic material embedded in the document including making them available to other applications as if each were a drawing pane in an arbitrary window. This enables them to be used to support visual activity ranging from bit map and line graphics, QuickTime videos, through simulation and animation, to

Embedded active concept maps can be used to support semantic networks (Gaines, 1991a) providing formal knowledge representation that can be compiled and run within the knowledge representation and inference system. Figure 8 shows a screen dump from KWrite editing a paper that was published in the proceedings of the British Computer Society Expert Systems Conference in December 1992 (Gaines and Shaw, 1992). The paper puts into active document form a knowledge-based system developed as a solution to a 'challenge' problem circulated by Project Sisyphus, an initiative of the European Knowledge Acquisition Workshop. The problem is one of room allocation derived from an ESPRIT project (Voß, Karbach, Drouven, Lorek and Schuckey, 1990). The embedded knowledge structures provide a complete solution to the room allocation problem (Gaines, 1994c).

The electronic version of the paper was made available through anonymous FTP and as a CD-ROM. It demonstrated parallel publication of the paper version as a camera-ready copy, book copy and the electronic version, identical in appearance, as a full working demonstration of the problem solution. What is particularly significant is that there were no hidden data structures. The semantic networks in the paper were the complete knowledge structures and operated directly in the inference engine to solve the problem. Thus, the document was also the knowledge base and editing the knowledge structures within the document can change the ontologies, rules or data, and hence the outcome of inference.

The multimedia and popup menu linkage facilities remain available, and, for example, the user is shown at the bottom right of Figure 8 as having accessed the popup menu associated with a particular node. One of the facilities this makes available is access to annotation for that node, and, since the script has full operating system access, this may be within the same document, in another application on the machine, or in another application across the network, possibly in another country.



The abstract of the BCS paper describes the form of publication:

## 7 Conclusions

Increasing ease of access to the Internet is making it feasible for geographically dispersed communities to work closely together, coordinating their activities through electronic mail, digital document archives, and access to remote computing facilities. This article has reported on practical experience of knowledge management for a number of communities and projects.

It has given an overview of the technologies available, their accessibility, ease of use and impact on collaborative activities. It has discussed the practical problems that arise, the limitations of existing technologies, and how these may be overcome.

Examples have been given of projects encompassing a range of multimedia digital technologies from list servers, through World Wide Web document archives, production in a few days of CD-ROMs containing movies and digitized foils and documents giving a complete account of working meetings, to active documents that are themselves expert systems.

The combination of multi-media, World Wide Web and artificial intelligence tools now available provides the technological foundations for very powerful knowledge management systems.

## Acknowledgments

Financial assistance for this work has been made available by the Natural Sciences and Engineering Research Council of Canada.

## References

- Bell, D. (1973). **The Coming of Post-Industrial Society**. New York, Basic Books.
- Flores-Mendez, R.A. (1997). Java concept maps for the learning web. **Proceedings of ED-MEDIA'97 : World Conference on Educational Multimedia and Hypermedia**. Charlottesville, VA, Association for the Advancement of Computing in Education.
- Gaines, B.R. (1989). Social and cognitive processes in knowledge acquisition. **Knowledge Acquisition** 1(1) 251-280.
- Gaines, B.R. (1990). Knowledge support systems. **Knowledge-Based Systems** 3(3) 192-203.
- Gaines, B.R. (1991a). An interactive visual language for term subsumption visual languages. **IJCAI'91: Proceedings of the Twelfth International Joint Conference on Artificial Intelligence**. pp.817-823. San Mateo, California, Morgan Kaufmann.
- Gaines, B.R. (1991b). Modeling and forecasting the information sciences. **Information Sciences** 57-58 3-22.
- Gaines, B.R. (1993). An agenda for digital journals: the socio-technical infrastructure of knowledge dissemination. **Journal of Organizational Computing** 3(2) 135-193.
- Gaines, B.R. (1994a). Class library implementation of an open architecture knowledge support system. **International Journal Human-Computer Studies** 41(1/2) 59-107.
- Gaines, B.R. (1994b). The collective stance in modeling expertise in individuals and organizations. **International Journal of Expert Systems** 7(1) 21-51.
- Gaines, B.R. (1994c). A situated classification solution of a resource allocation task represented in a visual language. **International Journal Human-Computer Studies** 40(2) 243-271.
- Gaines, B.R., Norrie, D.H. and Lapsley, A.Z. (1995). Mediator: an Intelligent Information System Supporting the Virtual Manufacturing Enterprise. **Proceedings of 1995 IEEE International Conference on Systems, Man and Cybernetics**. pp.964-969. New York, IEEE.
- Gaines, B.R. and Shaw, M.L.G. (1992). Documents as expert systems. Bramer, M.A. and Milne, R.W., Ed. **Research and Development in Expert Systems IX. Proceedings of British Computer Society Expert Systems Conference**. pp.331-349. Cambridge, UK, Cambridge University Press.
- Gaines, B.R. and Shaw, M.L.G. (1993a). Basing knowledge acquisition tools in personal construct psychology. **Knowledge Engineering Review** 8(1) 49-85.
- Gaines, B.R. and Shaw, M.L.G. (1993b). Eliciting knowledge and transferring it effectively to a knowledge-based systems. **IEEE Transactions on Knowledge and Data Engineering** 5(1) 4-14.
- Gaines, B.R. and Shaw, M.L.G. (1993c). Open architecture multimedia documents. **Proceedings of ACM Multimedia 93**. pp.137-146.
- Gaines, B.R. and Shaw, M.L.G. (1994a). Knowledge acquisition and representation techniques in scholarly communication. **Proceedings SIGDOC'94: ACM 12th Annual International Conference on Systems Documentation**. pp.251-260. New York, ACM.
- Gaines, B.R. and Shaw, M.L.G. (1994b). Using knowledge acquisition and representation tools to support scientific communities. **AAAI'94: Proceedings of the Twelfth National Conference on Artificial Intelligence**. pp.707-714. Menlo Park, California, AAAI Press/MIT Press.
- Gaines, B.R. and Shaw, M.L.G. (1995a). Concept maps as hypermedia components. **International Journal Human-Computer Studies** 43(3) 323-361.
- Gaines, B.R. and Shaw, M.L.G. (1995b). WebMap: concept mapping on the web. **World Wide Web Journal** 1(1) 171-183.
- GNOSIS (1993). TW4 Soft Machinery Workshop Proceedings. Knowledge Science Institute, University of Calgary, Canada. [ftp://ksi.cpsc.ucalgary.ca/IMS/GNOSIS/TC7/tw4ws2](http://ksi.cpsc.ucalgary.ca/IMS/GNOSIS/TC7/tw4ws2).
- GNOSIS (1994). Knowledge Systematization: Configuration Systems for Design and Manufacturing: Final Report of the Test Case. Knowledge Science Institute, University of Calgary, Canada.

- <http://ksi.cpsc.ucalgary.ca/IMS/GNOSIS> ,  
<ftp://ksi.cpsc.ucalgary.ca/IMS/GNOSIS/TC7/final>.
- Habermas, J. (1981). **The Theory of Communicative Action: Reason and the Rationalization of Society**. Boston, Beacon Press.
- Kremer, R. and Gaines, B.R. (1994). Groupware concept mapping techniques. **Proceedings SIGDOC'94: ACM 12th Annual International Conference on Systems Documentation**. pp.156-165. New York, ACM.
- Manitoba (1993). **Proceedings of the 1993 International Conference on Refereed Electronic Journals**. Winnipeg, Manitoba, University of Manitoba Library.
- Renear, A.H. and Bilder, G. (1993). Two theses about the new scholarly communication. Landow, G.P. and Delany, P., Ed. **The Digital Word: Text-based Computing in the Humanities**. pp.217-236. Cambridge, Massachusetts, MIT Press.
- Shaw, M.L.G. and Gaines, B.R. (1994). Active documents combining multimedia and expert systems. **Proceedings SIGDOC'94: ACM 12th Annual International Conference on Systems Documentation**. pp.166-175. New York, ACM.
- Voß, A., Karbach, W., Drouven, U., Lorek, D. and Schuckey, R. (1990). Operationalization of a synthetic problem. **ESPRIT Basic Research Project P3178 REFLECT Task I.2.1 Report**. Bonn, Germany, GMD.