

DAITC

AD-A209 318

HYPERMEDIA LABORATORY
DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER
REVIEW FOR 1988

DAITC/TR-88/014

Allan D. Kuhn
Director

December 1988



Defense Applied Information Technology Center
Hypermedia Laboratory

1800 North Beauregard Street
Alexandria, Virginia 22311
(703) 998-4787
Fax No. (703) 931-3968

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED		1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited	
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE			
4. PERFORMING ORGANIZATION REPORT NUMBER(S) DAITC/TR-88/014		5. MONITORING ORGANIZATION REPORT NUMBER(S) DTIC/TR/89/12	
6a. NAME OF PERFORMING ORGANIZATION Defense Applied Information Technology Center	6b. OFFICE SYMBOL (if applicable)	7a. NAME OF MONITORING ORGANIZATION Defense Technical Information Center	
6c. ADDRESS (City, State, and ZIP Code) 1800 N. Beauregard Street Alexandria, VA 22311		7b. ADDRESS (City, State, and ZIP Code) Cameron Station Alexandria, VA 22304-6145	
8a. NAME OF FUNDING / SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (if applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER	
8c. ADDRESS (City, State, and ZIP Code)		10. SOURCE OF FUNDING NUMBERS	
		PROGRAM ELEMENT NO.	PROJECT NO.
		TASK NO.	WORK UNIT ACCESSION NO.
11. TITLE (Include Security Classification) Hypermedia Laboratory, Defense Applied Information Technology Center; Review for 1988			
12. PERSONAL AUTHOR(S) Allan D. Kuhn, Duc T. Tran, Randy L. Bixby, Brenda S. Edleson			
13a. TYPE OF REPORT	13b. TIME COVERED FROM _____ TO _____	14. DATE OF REPORT (Year, Month, Day) 8812	15. PAGE COUNT 140
16. SUPPLEMENTARY NOTATION			
17. COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	Hypermedia, Artificial Intelligence, Bitmap Systems, Networking, Media, Gateways, Interfaces, Defense Applied Information Technology Center, DAITC, (Continued on reverse)	
5	2		
12	5		
19. ABSTRACT (Continue on reverse if necessary and identify by block number)			
<p>This review documents the activities of the Hypermedia Laboratory for 1988. The Laboratory began its existence with the incorporation of artificial intelligence into the DTIC Common Command Language System Design project for the DoD Gateway Information System. With the availability of bitmap systems, rapid prototype developments were done making use of these systems' capabilities of icon-drive, multi-windowing and multi-tasking. It was a short step into our expanding exploration of technologies involving integration of artificial intelligence, bitmap systems, networking, and electronic media as means of transmitting and storing information. In hypermedia, artificial intelligence will serve as the interface engine between the user and the information source; the bitmap system will be the platform for viewing information in a manner organized by the viewer; networking gets the user to the diverse and multiple sources of information; media concerns the diverse forms of storing information, including online, optical disc, floppy disc, voice, electronic mail, et.al., and stimuli information such as sound, (continued on reverse)</p>			
20. DISTRIBUTION / AVAILABILITY OF ABSTRACT <input type="checkbox"/> UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT <input type="checkbox"/> DTIC USERS		21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL Allan D. Kuhn		22b. TELEPHONE (Include Area Code) (703) 998-4600	22c. OFFICE SYMBOL

Block 18. Subject Terms (continued)

Defense Technical Information Center, DTIC, DoD Gateway Information System, DGIS, Scientific and Technical Information Network, STINET.

Block 19. Abstract (continued)

images, and color. The integration of these technologies will give the user the capability to collect and handle information, from any source, in a contextual manner. Context handling is the reason for implementing hypermedia systems. The integration of technologies for context handling is the primary goal of the Hypermedia Laboratory.

HYPERMEDIA LABORATORY
DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER
REVIEW FOR 1988

Allan D. Kuhn
Director

December 1988

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input checked="" type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	<input type="checkbox"/>
By _____	
Distribution/ _____	
Availability Codes	
Avail and/or	
Dist	Special
A-1	

2
QUALITY
INSPECTED

DRITC Hypermedia Laboratory Report Number 3

The Hypermedia Laboratory

Defense Applied Information Technology Center
Alexandria, VA 22312

Defense Technical Information Center
Alexandria, VA 22304

This report was prepared in the
DTICSP0 Hypermedia Laboratory

CONTENTS

PREFATORY

The Hypermedia Laboratory, Defense Applied Information Technology Center (DAITC). *Allan D. Kuhn*

Hypermedia: Our Experience in the Intertwining Zone. *Allan Kuhn*

Hypermedia Lab. Program Note 1, May 1988: Welcome to The DAITC Hypermedia Laboratory

Hypermedia Lab. Program Note 2, May 1988: Welcome to Hypermedia

BITMAP SYSTEMS

A Hypermedia System for DGIS: A First Technical View. *Duc T. Tran*

Notes on X WINDOWS Taken from the X WINDOWS Conference, Massachusetts Institute of Technology, January 1989. *Duc Tran and Jon Krueger*

Hypermedia Lab. Technical Note 1, Aug 88: The Symbolics Processor

ARTIFICIAL INTELLIGENCE

Toward an Artificial Intelligence Environment for Defense Technical Information Center (DTIC): Proposed Tasks -- 1987. *Allan D. Kuhn*

The DGIS Common Command Language System (CCLS): DGIS Steps into Artificial Intelligence. *Allan Kuhn*

Enhancements Outline for the DoD Gateway Information System (DGIS): Proposal for AI/ES-Based Information Modelling Technology. *Richard S. Marcus*

Hypermedia Lab. Program Note 3, May 1988: DGIS Common Command Language System Project & Artificial Intelligence

Hypermedia Lab. Technical Note 2, Aug 88: Object-Oriented Perspective

Hypermedia Lab. Technical Note 3, Oct 88: Genetic Algorithms

Hypermedia Lab. Technical Note 4, Nov 88: Connectionist Models (or Neural Networks)

Hypermedia Lab. Technical Note 5, Dec 88: C Language Integrated Production System (CLIPS)

Hypermedia Lab. Technical Note 6, Dec 88: PROLOG -- An Artificial Intelligence Programming Tool

Hypermedia Lab. Technical Note 7, Dec 88: Quintus PROLOG: A Vendor Package Version

Hypermedia Lab. Technical Note 8, Dec 88: NU-PROLOG: An Example of a University Version

NETWORKING

Global Scientific and Technical Information Network. *Gladys A. Cotter*

Possibilities Actuelles et Futures de la Technologie / Information Technology: Current and Future Possibilities. *Albert Yanez*

A Look at Some of the Elements of Translation: Preparatory to Accessing Global Information. *Allan Kuhn*

Networking and Interoperability Laboratory, Defense Applied Information Technology Center. *Judy F. Hunter*

DAITC Network Expert System. *Brenda A. Edelson*

Hypermedia Lab. Technical Note 9, Dec 88: X-WINDOWS

MEDIA

Hypermedia in DoD: Where The Medium Is The Message. *Brenda Edelson*

Hypertext in a Resource Sharing Greater Machine Level Environment. *Randy L. Bixby*

Hypermedia Lab. Program Note 5, October 1988: Hypermedia & Hypertext

Hypermedia Lab. Program Note 6, October 1988: Voice & the Gateway Computer System

Hypermedia Lab. Program Note 8, December 1988: minicomputer - Hypertext - multiuser

SUMMARY

Summary & A Presentational Overview of the DAITC Hypermedia Laboratory Activities. *Allan Kuhn*

Hypermedia Laboratory Goals for 1989.

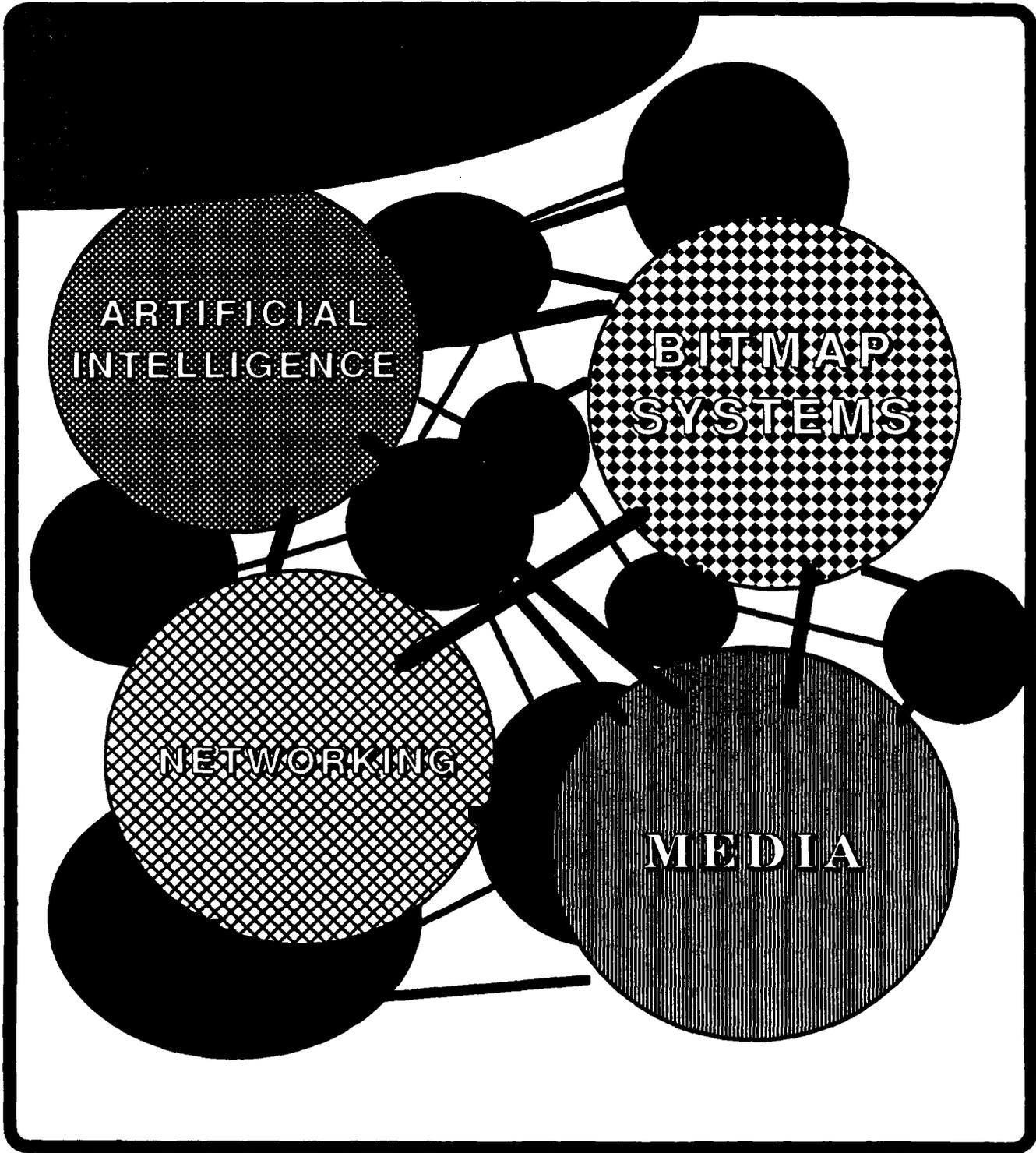
APPENDICES

Hypermedia Lab. Program Note 7, Dec 88: Presentations & Publications 1988

Hypermedia Lab. Program Note 4, Aug 88: Forerunner Publications 1987

Personal Acknowledgements - 1988

HYPERMEDIA IS THE MERGING OF



**THE HYPERMEDIA LABORATORY
DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER (DAITC)**

*Allan D. Kuhn
Director, Hypermedia Laboratory*

In the most general terms, the Hypermedia Laboratory of the Defense Applied Information Technology Center is a place for the integration and implementation of new information technologies, particularly those which touch people directly. More specifically, the Hypermedia Laboratory is one of the component laboratories of the DAITC, and constitutes one of the two DAITC technology integration laboratories. The other is the DAITC Networking and Interoperability Laboratory. The Hypermedia Laboratory concerns itself with integrating the individual technologies that are merging for presenting, or displaying, the information, whereas the Networking and Interoperability Laboratory concerns itself with integrating network and machine system technologies for transferring and handling that information towards its presentation.

The prime goal of the Hypermedia Laboratory is to create artificial intelligence- and hypermedia-based programs that make the human-machine interface more human in the machine's responses.

In addition to the Hypermedia Laboratory's internal development activities, examples of individual technologies for which there are laboratories at the DAITC are High Density Information Storage (e.g., CD-ROM), Interactive Media (video-disk), Voice, Text Search (Hypertext), Gateway, *et al.* The Hypermedia Laboratory works with those laboratories in cooperative developments that integrate their technologies into Hypermedia systems.

Hypermedia offers new possibilities for accessing large and complex information sources. Employing the emerging hypermedia concepts, information from those sources can be organized in different ways, depending on individual viewpoints. Hypermedia allows the user to organize database information in a visual way by using the capabilities of bit-mapped systems, such as icons and multiple windowing techniques. This visual representation lets the user adapt the data to one's needs in greater or less detail, depending on one's level of knowledge or expertise. Hypermedia allows the user to organize text, pictures, and sound by association, by context, the same way humans organize information in their minds. Contextual and spatial cues supplement the user's model of which nodes, or points of reference, one is viewing and how they are related to each other and their neighbors in the displays. Hypermedia is the window on the worlds of text, video, and sound. It is an information medium that bridges the gaps between the existing media, provides us with ways of coping with massive amounts of information, and gives us an easy way to learn and use interfaces for accessing heterogeneous databases.(1)

Internally, the Hypermedia Laboratory has a Symbolics 3650 Processor that includes a good range of software packages, including an Intellicorp KEE Expert System Building Shell, a PROLOG package for integrating PROLOG- based programs, of which there are currently several on the DoD Gateway Information System (DGIS) and other (see Hypermedia Lab Technical Note 1 for fuller description). There are also two SUN 3 workstations (160 and 260), and two XEROX 1186s. The Laboratory systems, except for the XEROXs, are netted to the DAITC computer systems. Because of the DAITC network, the Hypermedia Laboratory has access to systems throughout the DAITC outside of the Laboratory. The Laboratory also has a Quintus PROLOG compiler/interpreter/debugger package resident on the DAITC/DGIS VAX 11/780; this package is a high quality software item with numerous capabilities and good maintenance support.

The Laboratory serves DoD through funded projects that will benefit the whole of DoD through exportable developments. The Hypermedia Laboratory has operated as an entity since April, 1988, and is staffed by three technical information science experts, two at MS level and one at PhD level, with extensive

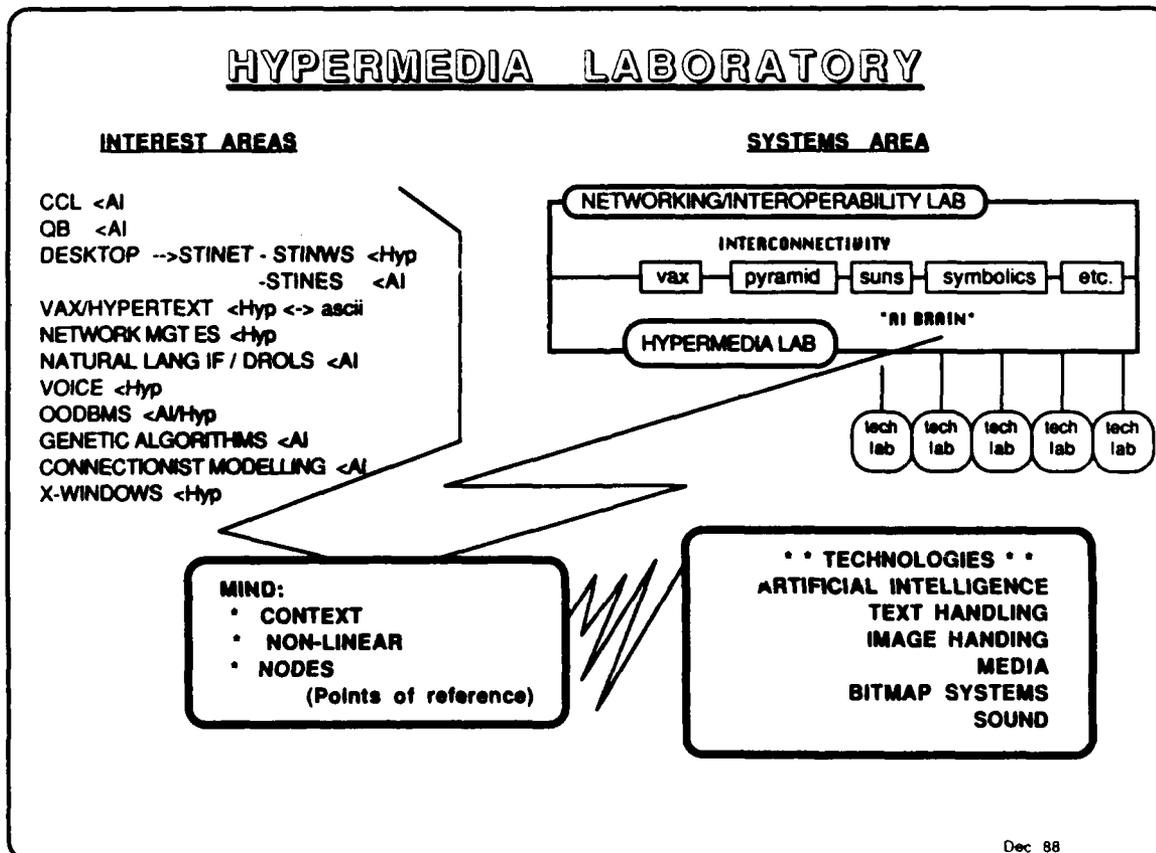
knowledge in requirements for handling information from source to enduser, and two computer science experts, one at BS level and one at MS level, both with tracks in artificial intelligence. The Laboratory furthermore freely draws on the multiplicity of expertises available and associated with the other DAITC areas in an *across-the-walls* environment.

The Hypermedia Laboratory and the whole of the DAITC approach is to provide responsiveness through rapid project startups, practicality through workable solutions to information handling problems, cost-efficiency through low investment for prototype startup and sharing of development results, and Government direction supported by agency-to-agency agreements.

DAITC Hypermedia Laboratory Director:

Allan D. Kuhn,
DAITC: (703) 998-4600
DAITC Office: (703) 998-3451
DTIC-EB, (202) 274-5367, AV 287-5367
Electronic Mail: daic!kuhn

(1) This description for Hypermedia was formulated by Randy L. Bixby, Hypermedia Laboratory, and Judy F. Hunter, Director, DAITC Networking and Interoperability Laboratory, based on other sources; the idea is to have a universally acceptable definition and description of the concept of Hypermedia.

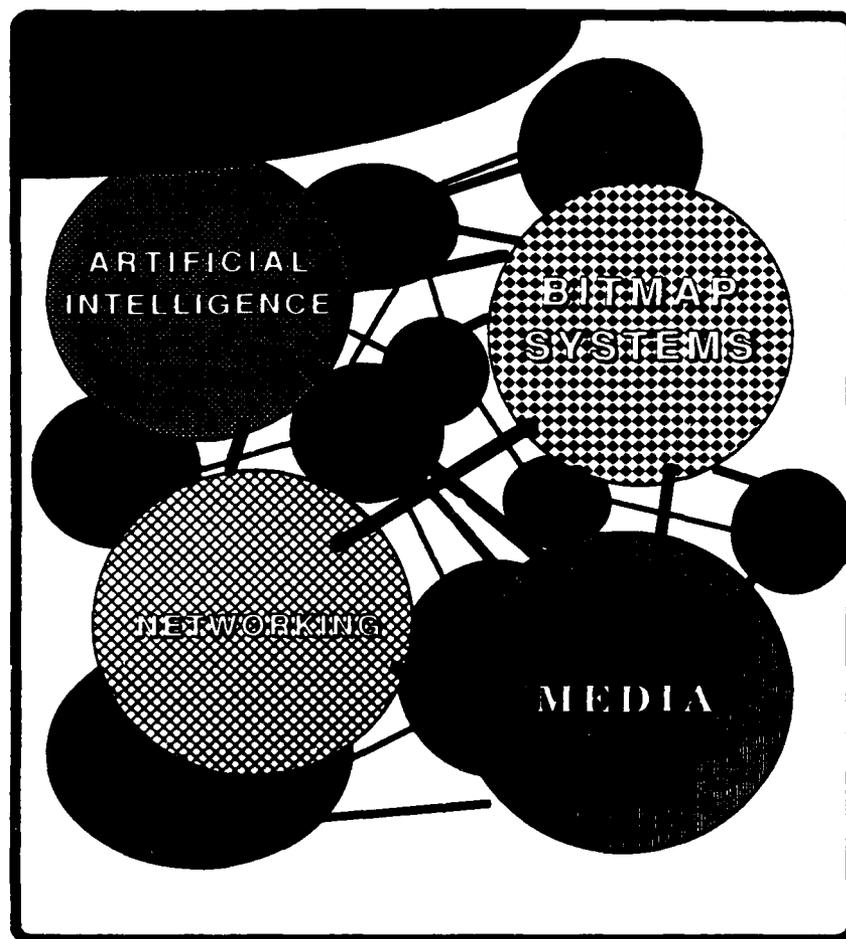


HYPERMEDIA

OUR EXPERIENCE IN THE INTERTWINGLING ZONE

*Allan Kuhn
Hypermedia Laboratory*

In our view, hypermedia is the merging of technologies and media information sources through



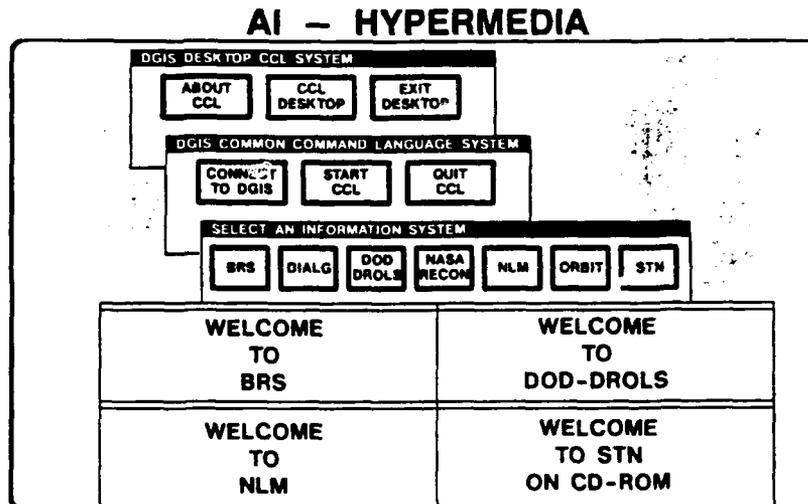
The convergence and integration of these technologies allow the user to acquire and organize comprehensive information, including stimuli information such as sound, color, and images, in a nonlinear manner that begins to approach the way the mind organizes information. Hypermedia implementation is creating a dramatic change in the electronic handling and processing of information. As an expression of these hypermedia concepts, the melange of Hypermedia Laboratory interests includes desktop environments, object-oriented environments, hypertext, voice, knowledge base and blackboard architectures, and the emerging genetic algorithm environment with its ramifications for involving connectionist modelling.

Bitmap systems are becoming popular and are proliferating throughout DoD. These systems encompass

not only workstation-size systems but also desktop size units, which are the "new" personal computers. These systems, regardless of their physical size, have capabilities that now permit the integration of media, and therefore open the door to hypermedia implementations.

Imagine sitting down at your workstation or personal computer, and initiating simultaneous displays of information from an online database, a CD-ROM database, a file that you have created yourself that includes figures and graphs, and a high density video disk that is a general comprehensive reference such as an encyclopedia. Further imagine having this set of displays include color, text information, static images, motion imagery, and music. Verbal information that supports the visual information. On top of all this, all interactions and interfaces have been invoked and are controlled through icons, some of which are programmed into the central DoD Gateway Information System (DGIS), some of which come with media peripherals (such as the CD-ROM on its player), and some of which you create yourself on your workstation.

Our entry into hypermedia began with looking at a bitmap system desktop environment for the DGIS Common Command Language System (CCLS), which is PROLOG-based. With CCLS as the vehicle, we found ourselves integrating bitmap system capabilities with a search assistant that is driven with an AI engine.



Using a SUN 3/260 workstation, we made use of icons, windowing, simultaneous database search session displays, and color. We foresaw also a hypermedia link to CD-ROM information bases as an inclusion in this hypermedia system; of immediate application to the DoD information community is the prototype DTIC DROLS Technical Report Database on CD-ROM.

We feel that getting into the hypermedia environment and looking at the multitude of possibilities will keep us ahead of what is already coming down the road in the DoD community: bitmap systems and hypermedia peripherals in the enduser environment. What we see taking place immediately are:

- o Icon drive - creating iconic systems to invoke DGIS operations.
- o Multiple window displays - using the windowing technology of the bitmap system.
- o Simultaneous tasking - parallel and disparate tasks operating concurrently.
- o The use of color - a stimulus medium which is a part of the information message.
- o Multimedia accesses - online, CD-ROM, video disc, other storage media.
- o Intermedia accessing - simultaneous access of disparate media.
- o Simultaneous disparate window displays - text, static images, motion.
- o Sound - verbal and musical information supporting the visual information.

Our goal is to create an integrated hypermedia linkage that serves as an adjunct to DGIS, does not replace ASCII-based DGIS, but co-exists with it as one of its power tools for the users who desire the capabilities.

Hypermedia is the convergence, therefore, of information technologies which, in their interfacing, allows you to organize text, pictures, and sound by association or by context on a machine system. Hypermedia offers new possibilities for accessing large or complex information sources. The already well-developed bitmap system technology allows the user to organize information in a visual manner through simultaneous window displays. This visual representation lets the user adapt the data to what he needs in greater or lesser detail depending on his level of knowledge or expertise. Hypermedia allows the user to organize text, pictures, and sound by association or by context the same way humans organize information in their minds. Contextual and spatial cues supplement the user's model of which nodes he is viewing and how they are related to each other and their neighbors in that model. Hypermedia is the window on the worlds of text, video, and sound. It is an information technology that will bridge the gaps between the existing media and provide ways of coping with massive amounts of information.

alpha

The primary vehicle for entry into hypermedia applications and implementation is the **bitmap system** with its graphics, multiwindowing, and multitasking capabilities. The lowering cost of these systems (or workstations), resulting from the ever-evolving development in the technology, is making these systems just as affordable for the DoD technical library and enduser communities as traditional character-based microcomputer and workstation systems. The concurrent technological developments in storing and accessing information in various media, such as floppy disks, CD-ROM, video disks, in addition to online systems, are the other components that create the hypermedia system. The popularity of bitmap systems is rapidly expanding in the DoD enduser community, such as office staffs, researchers, and engineers; information in all its possible formats (text, images, sound) stored in high-density media (CD-ROM, etc.) is now widely and easily available on the open market.

beta

An adjunct entry to hypermedia activity is **integration into networking**. Networking ranges from local area internal networks to accessing wide area external networks having wide geographical communication. These networks and their communications links in turn range from village to global. Networking provides access to and retrieval from the plethora of both internal and external sources of information. With the rapid development of technologies, we are not only talking of accessing character-based systems, but also remote graphics information and programs. This will be possible with the emerging *X-Windows* development coming out of the Athena Project at the Massachusetts Institute of Technology.

gamma

Because of the complexities of networking, accessing, and retrieving from the vast multiplicity of information sources, **artificial intelligence** serves as the engine that drives the interfaces between the user and the information source. DTIC's first step into this area is the development of a Common Command Language System which is PROLOG-based. The purpose of this AI interface is to provide a universality in searching all systems, including in a manner that helps the searcher to search unfamiliar systems. The dream of developers in incorporating AI is to create machine systems that communicate with humans on an intelligible level.

delta

Media serves as the multiplicity of sources of diverse information. We now can access online information systems, CD-ROM information bases, and information on other types of optical disks such as WORM (Write Once, Ready Many), CD-Audio, analog video disks, CD-Interactive, and the emerging DVI (Digital Video Interactive). The controversial digital audio tape (DAT) will become a media resource.

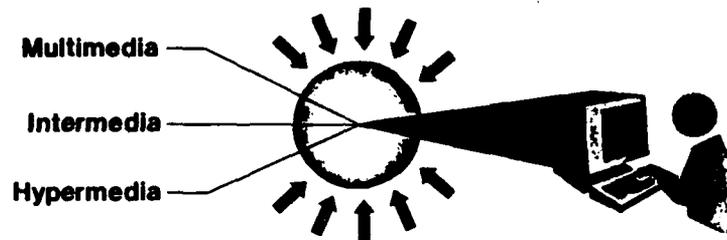
Furthermore, *any* electronic source serves as a source of information; floppy disks, electronic mail, digitized voice, video conferencing, *usw.* Media also involves the human senses, based on cognition, of hearing (speech, music, cacaphony), vision (color, still and moving images, *trompe-l'oeil*), and touch and smell are being worked on. Media is further matrixed through *multimedia*, *intermedia*, and *hypermedia*. Each is a progression to the other in terms of branching, linking, integrating. I would suggest that:

o **Multimedia** -- is the body of media, whether electronic or paper, each medium treated and handled separately.

o **Intermedia** -- recognizes that information is stored in a range of disparate media, that each medium should have access established to it in order to acquire a comprehensive set of information, and that no medium should be excluded from accessing because of its type of medium. Paper media, with its *text* and graphics, can be scanned into digitized form.

o **Hypermedia** -- encompasses the range of information in view of human cognition, and its purpose is to allow the user to arrange a selected set of information views in a contextual manner.

TECHNOLOGY CONVERGENCE



The Defense Technical Information Center (DTIC), in sponsoring and working through the Defense Applied Information Technology Center and the DoD Gateway Information System, is at its initial stage in its entry into hypermedia. Because the business of DTIC is information, there is a vast opportunity for looking at hypermedia applications, assessing their relevance to the handling and processing of information in the DoD community, and developing hypermedia-based implementations for doing so. Additionally, as DTIC explores this area more intensively, we find a much greater requirement to coordinate and interface all developments done by DTIC to aid the DoD user in the quest for information.

If DTIC is to take the lead in establishing a unified Scientific and Technical Information (STI) network for the DoD information worker, there are several coalescing agents needed to create such a network system, all having equal importance. For example, one is interoperability, another is the identification of user needs. Another such agent is the capability to collect and handle information, from all sources, in a contextual manner. Context handling is the reason for implementing hypermedia systems.

Extract Reference; CCL in Hypermedia:

Kuhn, Allan D., and Gladys A. Cotter. "The DoD Gateway Information System (DGIS): The Department of Defense Microcomputer User's Gateway to the World;" *Microcomputers for Information Management: An International Journal for Library and Information Services*, v5 n2, June 1988, pp.73-92

Reading References:

Global Scientific and Technical Information Network.

Gladys A. Cotter, Defense Applied Information Technology Center, December 1988, AD-A201 902. Also published in Proceedings, ONLINE INFORMATION 88, 12th International Online Information Meeting, London, 6-8 December 1988, pp. 611-618.

Discusses technologies for worldwide networking and information accessing.

"Knowledge Gateways: The Building Blocks." *Information Processing & Management*, Volume 24, Number 4, pp. 459-468, 1988.

Donald T. Hawkins, Louise R. Levy, AT&T Bell Laboratories, and K. Leon Montgomery, School of Library and Information Sciences, University of Pittsburgh.

Discusses the technologies required in getting the user to the multiplicity of information systems; corresponds to the needs of hypermedia.

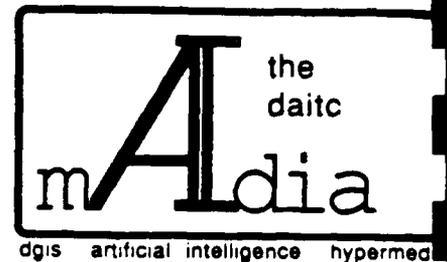
Networking and Interoperability Laboratory.

Judy F. Hunter, Defense Applied Information Technology Center, [November 1988].

Discusses the DAITC areas of interest in internal and external networking.

**EVERYTHING IS DEEPLY INTERTWINGLED.
Especially in the well-designed interactive system.**

Ted Nelson, DREAM MACHINES (p.70), the flip side of COMPUTER LIB, Tempus Books, 1974,1987.



Welcome to
* * * THE DAITC HYPERMEDIA LABORATORY * * *

What is the Hypermedia Laboratory?

The "Hypermedia Laboratory" is the designation of the activity at the Defense Applied Information Technology Center (DAITC) that is concerned with the technologies convergence taking place in artificial intelligence and hypermedia. This activity is managed by the Defense Technical Information Center (DTIC). Much of our development at this point is mainframe-oriented in support of the DoD Gateway Information System (DGIS). We expect these developments to be exportable, however, not only to other mainframe environments, but also to workstation and PC environments.

The basic thrust of the term "hypermedia" is accessing and transferring any kind of information available in any kind of medium, making use of the human senses of comprehension, vision, and hearing. Although the designation "hypermedia" is relatively new, its concepts are not. A major center intensively involved in advanced interdisciplinary AI-Hypermedia studies in eleven major research areas is the Media Laboratory, Massachusetts Institute of Technology, operating since 1975.

The constitution of our Hypermedia Lab. currently includes a Symbolics 3650 processor and a SUN 3/260 workstation, in addition to the facilities of the DGIS and DAITC computer systems which run on UNIX operating systems. The DGIS VAX includes a PROLOG compiler and interpreter package, which initiated our stepping into artificial intelligence about one year ago. All these systems are internettted so that in communicating with each other, we are in effect creating a virtual large AI "brain" as a component of the DGIS system. We have on order two MAC II systems of mixed application configurations. Although the Symbolics and the SUN, in addition to their AI capabilities, put us across the threshold of hypermedia, the MAC II systems will open up greater hypermedia capabilities and entrench us even more in this area.

Hardware/software is only one essential element of the Lab. The critical element is expertise. We are highly fortunate in having available a multitude of DAITC people having extensive technical skills and expertise, and who are highly motivated. The Hypermedia Lab. staff itself is made up of two technical AI people, one a senior technician having extensive knowledge in AI application and implementation, the second having a strong LISP background and skills in Expert System construction. The staff also includes two DTIC people, one acting as the Lab. program manager, the second a technical information/ library science domain expert.

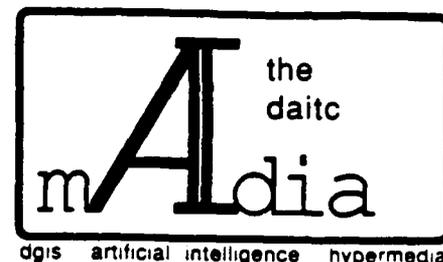
The Hypermedia Lab. works closely with the DAITC Networking/Interoperability Laboratory. In doing so, we have access to the expertise of the DAITC systems people. This is mutually beneficial in that all concerned have access to all DAITC developments and the interfacing of those developments. The AI "brain" is dependent on the interest and participation of the systems people, who in turn contribute to the enhancement and expansion of the Hypermedia Lab. activities.

The Lab. is also fortunate in being strongly supported by expert management provided by the DAITC management; this is not only the DTIC management of the DAITC program, but also the DAITC facility management provided by TRESP.

For a brief on "Hypermedia," see Hypermedia Lab. Program Note 2.

The prime goal of the Hypermedia Lab. is to create AI-based programs that make the human-machine interface more human in the machine's responses.

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER (DAITC)
ALEXANDRIA, VIRGINIA



Welcome to
>>>>>>>>> HYPERMEDIA <<<<<<<<<<<

Imagine sitting down at your system, and initiating simultaneous displays of information from an online database, a CD-ROM database, a file that you have built yourself that includes figures and graphs, and a high density video disk that is a general comprehensive reference such as an encyclopedia. Further imagine having this set of displays include color, text information, static images, motion imagery, and verbal or music information that supports the visual information. On top of this, all interactions and interfaces have been invoked and are controlled through icons, some of which are programmed into the DGIS, some which come with the media peripherals and some of which you have created yourself.

Far out? Guess what -- it's already here. Systems are in place that include the processing of stimulus information as well as intellectual information as a part of the information transfer process.

HYPERMEDIA is the convergence of information technologies which, in their interfacing, allows you to organize text, pictures, and sound by association or by context on a machine system. This approaches the manner in which the mind organizes information. The human senses currently involved concentrate primarily on visual and and sound perceptions, along with the human sense of understanding. Other sensory applications, however, are being developed. The MIT Media Laboratory is working on automated touch; this is to be used as an information base of the viscosities of oils. A group at the University of North Texas is working on automating smell; using neuro-electronic assemblies, oderants are flushed across the assemblies, with electrodes shunting the signals into a computer that helps interpret the neural responses.

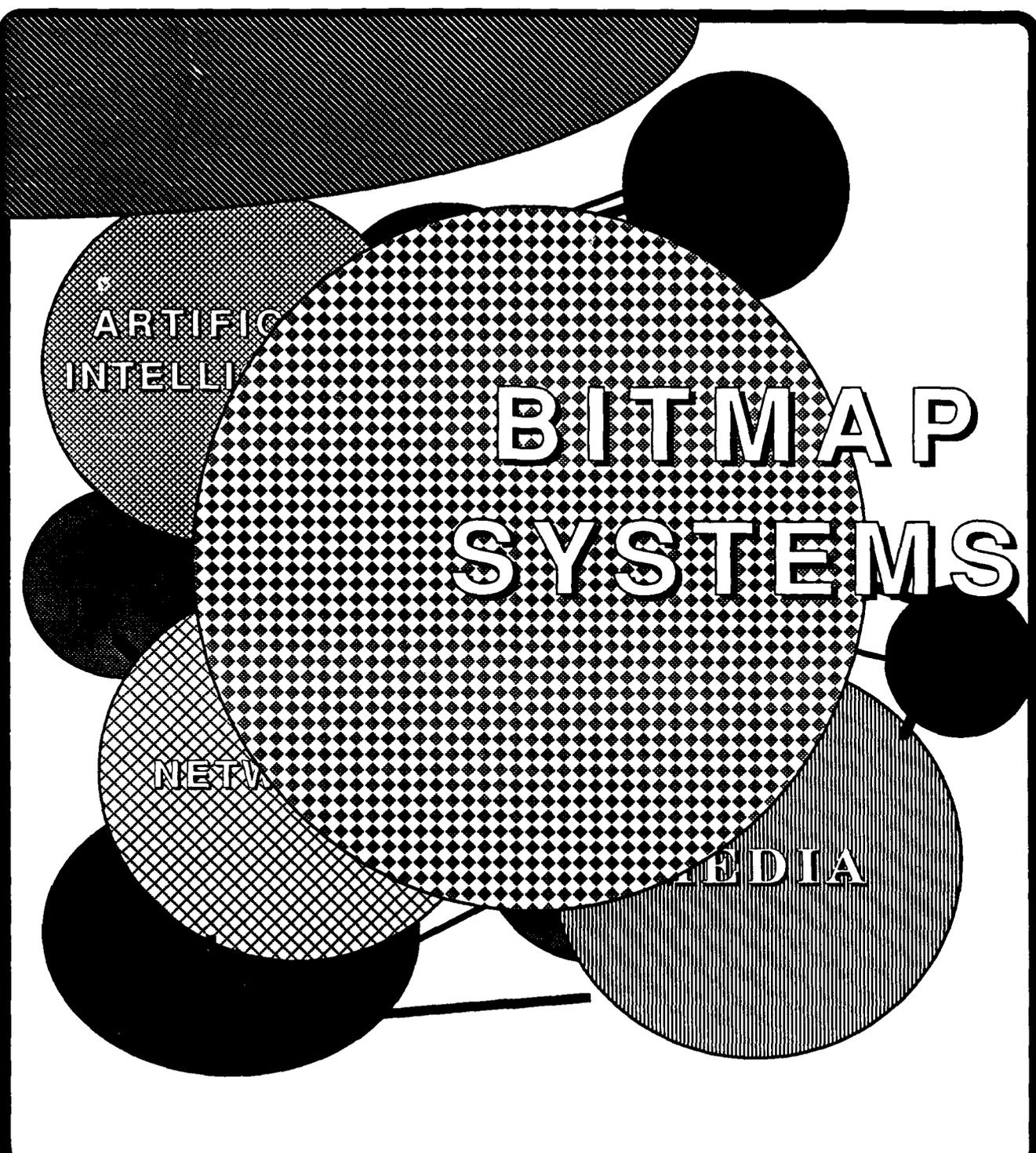
The hypermedia interests of the Hypermedia Lab. are based on supporting DGIS operations for DoD users having bitmap systems, such as MAC IIs and SUN workstations. Use of bitmap systems is rapidly expanding in DoD, e.g., note the expanding popularity of the MACs, and the 4,000 SUN workstation EDS contract for DoD. Another aspect is the incorporation of Symbolic Processor systems as frontends to existant databases.

The Lab. is currently exploring a desktop environment for the DGIS Common Command Language, as our initial entry into hypermedia. This means making use of icons, windowing, and simultaneous viewing of multiple database accesses. Color is an enhancement. This is being done on a SUN workstation, integrating UNIX, PROLOG-based programs, and bitmap system capabilities. We are also projecting a hypermedia linkage to a CD-ROM database, for example, DROLS.

We feel that getting into the hypermedia environment and looking at the multitude of possibilities for integrating media will keep us ahead of what is already coming down the road: bitmap systems and media peripherals in the enduser environment.

The prime goal of the Hypermedia Lab. is to create AI-based programs that make the human-machine interface more human in the machine's responses.

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER (DAITC)
ALEXANDRIA, VIRGINIA

An abstract graphic design featuring several overlapping circles and a large central circle. The circles are filled with different patterns: a fine grid, a coarse grid, a vertical line pattern, and solid black. A large, dark, curved shape is at the top left. The central circle is filled with a dense grid pattern and contains the text 'BITMAP SYSTEMS'.

ARTIFICIAL
INTELLIGENCE

BITMAP SYSTEMS

NETWORK

MEDIA

DAITC HYPERMEDIA LABORATORY

A HYPERMEDIA SYSTEM FOR DGIS

A First Technical View

Duc T. Tran
Hypermedia Laboratory

April 1988

1. Introduction

The DoD Gateway Information System (DGIS) as it currently stands provides users with the capability to access ASCII-based text information in remote databases, and download the information into local files. The medium for storing such information is text files. One may then process those files with supplied tools, *i.e.*, postprocessing routines. New information created by the postprocessing tools is also stored in the form of text files.

This view suggests that the text file format, purely ASCII, limits the flexibility of DGIS. Although bibliographic information is primarily text, once processed with appropriate tools it can be structured, enriched, and linked with information from other media, such as graphics, images, and sound. This linking with other media would put the DGIS into the realm of *hypermedia*.

The goal of a *hypermedia* system would be not to replace, but to coexist with ASCII- (or character-) based DGIS as one of its power tools for the users who desire the capabilities. The design below of such a system emphasizes affordability and simplicity in integrating existing hardwares and softwares.

The following describes how such a *hypermedia* system is possible. We present a software architecture of a *hypermedia* system in terms of its functional components, and then a hardware configuration that supports the design. A capability that is not emphasized here is the employment of artificial intelligence techniques. Our initial thrust is in *integration* to demonstrate the feasibility and usefulness of a *hypermedia* capability for DGIS. We feel that artificial intelligence will fit naturally into the system after proof of concept.

2. Software Architecture of a Hypermedia System

The DGIS hypermedia system consists of the following components:

2.1 The Hypermedia Integrator

The Hypermedia Integrator directs the control of the components of the system. It must have the capability to link, operate, and interface with different media. The linkages are often non-linear and non-structured in an intuitive manner so that the flow of control is not disrupted because of the absence, presence, and difference of components and media.

2.2 The Hypermedia Interface to DGIS

The purpose of this component is to communicate with DGIS to conduct searches in diverse databases, capture information in local text files, and process that information. This component will report the success or failure of its operations to the hypermedia system.

2.3 A DGIS Hypermedia Extended Toolbox

DGIS now has an ASCII-based toolbox for reformatting, or tailoring, ASCII text information. The purpose of this toolbox is to provide a number of standard tools to postprocess the captured text files. The basic

characteristics of these tools are standard batch processing with little or no user intervention, and no linkages to other media even when their availability is known. Examples of these tools are concordance generation and the hardly used graphical analysis (*i.e.*, the system-dependent trend analysis routine).

The Extended Toolbox is meant to be used in *ad hoc* interactive processing of text and non-text media. Cross-links or references to other media are either suggested or automatically established as availability is indicated. Here we are thinking about the capability for the user to link pieces of files, both text and non-text, to diverse media-based information sources such as a dictionary, an encyclopedia on CD-ROM, voice and sound in digitalized formats, data points on graphs to their detailed information, *etc.*

3. An Example of a Hardware/Software Hypermedia System Configuration

The Macintosh II system is an example of an available hypermedia system that could be incorporated into the DGIS environment. The MAC II capabilities include image scanning, sound recording and playing, high resolution bit-mapped graphics display, search tools for text and non-text information, and tools for composing and processing *hypermedia* information. The Hypermedia Controller is assumed by *HyperCard*, a multifaceted utility from Apple. The flow of the system is coded in the programming language of HyperCard, called *HyperTalk*. This language supports the non-linear approach to information management between different media. It can intuitively invoke the functions of each component by its friendly user interface that is typical of Macintosh.

Each component of a hypermedia-based system could be implemented as a Hypercard application. Typically, a Hypercard application is built as a set of interlinked *stacks*. Each stack is a logical part of the application. A stack consists of *cards*. A card presents information in terms of text and image. It also provides other action-oriented capabilities such as search, browse, and edit.

In this suggested configuration, graphics display and generation are supported by *Cricket Graph*, where a flow of frames can be controlled by *VideoWork*, an animation software package.

Large information bases could be accumulated and stored in a DBMS, the *4th Dimension*, for fast access and updating. Other large information bases could be made accessible to the Macintosh by the following methods:

- o CD-ROM disks read by an attached peripheral Apple CD-ROM player.
- o Interface with Relational DBMSs, such as ORACLE, INGRES, through an interface module.

Sophisticated functions of the system would be coded primarily in two languages: C provided by A/UX, and *Common LISP* by Coral software. Expert System capabilities, whenever applicable, would be implemented by two powerful expert system shells, the *NEXPERT* and *MacSmarts*.

4. Summation

In this case example, the importance of the birth of HyperCard, with its use of stacks, as a hypermedia integration utility has been recognized by Apple's users and developers. There is a vast market for making use of stacks, or *stackwares*. We have identified a number of stackwares that would be very useful to DGIS users. These stacks, generated by governmental sources and private industry, some in the public domain and others in the commercial market, could be provided as hypermedia *library stacks* with which the user would cross-link his information sources, or import them into his files. The "stacking" of one's information -- text, image, and even sound -- and to use and arrange that information in a manner that best suits the user's current purpose, is the creation and use of a **Hypermedia System**.

NOTES ON X WINDOWS TAKEN FROM THE X WINDOWS CONFERENCE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY, JANUARY 1989

Duc Tran
Hypermedia Laboratory

Jon Krueger
DoD Gateway Information Systems Laboratory

The Hypermedia Laboratory established an interest in the X WINDOWS development during 1988. These notes on the January 1989 Conference are included here as our first formal involvement in incorporating this technology into the Hypermedia Laboratory activities.

I.

From @dgis.daitc.mil:hvo@media-lab.media.mit.edu Tue Jan 24 01:47:10 1989
Subject: a few notes from the X windows conference

We would like to submit a few **NOTES AND OBSERVATIONS**
DIRECT FROM the campus of the Massachusetts Institute of Technology,
where the **X WINDOWS CONFERENCE** of 24-26 January 1989 is currently in progress.

1. **X WINDOWS on Personal Computers.**

We were most interested in obtaining evaluations of the current commercial products, and in learning what the state of the art is. We discovered that it is still far from off-the-shelf software, and not yet feasible. The problem is that a fully configured PC would still perform poorly for X WINDOW clients that used any graphics. It might be satisfactory to run a few concurrent terminal sessions.

2. **Xlib tutorial.**

This is the programmer's interface to X WINDOWS. It is a low-level interface, very powerful but very complex. It consists of a large library of function calls, currently difficult to master.

3. **X toolkit tutorial.**

This is a higher-level interface to program X WINDOWS. The level of abstraction achieved is very impressive. It compares favorably with the programming interface provided by Quintus PROLOG: it's an object-oriented approach which hides the complexity of Xlib.

4. **X WINDOWS and the Hypermedia Lab. Symbolics.**

We obtained enough information to install X on the Symbolics LISP machine. We are looking forward to the prospect of a common windows programming environment on the Pyramid, VAX, SUN, and the Symbolics LISP machines. This will allow us to write portable programs for distributed applications that implement high-quality user interfaces.

5. **Demonstration of the X terminal.**

This is explained by analogy --

terminal:timesharing::X terminal:workstation.

The advantage is that it makes X WINDOWS available to users without requiring them to buy a SUN and run UNIX.

6. X WINDOWS technology.

We observed that X WINDOWS has come of age as a mature technology. Like UNIX, it initially drew support from a large and enthusiastic user base. But as with UNIX, the availability of vendor support for hardware and software built on the X WINDOWS base has transformed the market. Of the hundreds of attendees, we found most were NOT hackers or even from colleges. They're about evenly mixed between developers looking to build X WINDOWS applications and systems integrators looking to buy them and resell them on their hardware.

7. This conference is mostly technical. Although vendors are present (and advertising!) it is aimed at developers. There's a commercial conference coming up on 26 June 1989 in San Jose.

II.

From hvo@media-lab.media.mit.edu Wed Jan 25 04:38:18 1989 Subject: tuesday at X conference

SEVERAL APPLICATIONS AND ISSUES FOR DAITC SYSTEMS

1. For sophisticated applications, we can program the Xlib for multiple overlay planes. These are like using several viewgraphs, laying them on top of each other. This is used to mix display features under user control, such as enabling and disabling tick marks on a map. You can obscure, overlap, or combine the features on each overlay plane. For instance, the user might flip between political and geographic information on a map, or display both at once. Or similarly, the user might turn on different groups of annotation on a map, or enable and disable display of levels of political boundaries, e.g. show state borders, show county borders, show each property line.

2. We saw the (unclassified) parts of an application for the Navy. As it happens, it doesn't use overlay planes, but its aims are similar. It displays a color map of any part of the world and shows where U.S. fleet components are located in that region. A novel and useful interface idea is their dynamic menu, which presents items based on options available and real world data from the map itself. For instance, if the user has already chosen a region, the item for choosing a region disappears. More interestingly, if there are only a few ships in that region it doesn't present the item soliciting which ships to display. Another interface idea provides ways to tailor the amount of annotation displayed. By default, the names and types of ships are displayed at their current locations. If the user wants information about the strength of each ship, one chooses an item from a pop-up menu, and the annotation is added.

3. We found that everyone wants to extend X. Extensions to handle 3D are being worked on; the X WINDOWS consortium has selected SUN Microsystems to implement them. Extensions to better handle X terminals, such as communications protocols better tuned for serial communications (say for an X terminal you use from home over a modem), are being evaluated. Extensions to handle color with better device independence so your hardcopy looks like your screen have been proposed by Tektronix.

4. Everyone's interested in display Postscript. This is because Postscript printers are widely available, and if Postscript imaging is supported by the window system, you can get a hardcopy without having to support device drivers for different hardcopy devices. DEC has developed a toolset that supports display Postscript in an X window. Their implementation is quite complete: if you could pipe the Postscript that Macintosh "MacDraw" generates for a Laserwriter to an X system, you could display the "MacDrawing" on your X terminal. Or you could generate the Postscript locally just as a convenient notation for drawing text and graphics. We observed this happening in real time; the speed was impressive.

5. Although the Xlib and X toolkit interfaces provide a great deal of data abstraction, LISP programmers still want an interface to functional languages. This is being done: CLX and CLUE are the LISP-based versions of Xlib and the X toolkit.

III.

Date: Wed Feb 1 08:38:31 1989 From: duc (Duc Tran) Subject: wednesday at X conference

EXAMPLES OF DEVELOPMENTS AND IMPLEMENTATIONS

1. SUN showed off their View2 Toolkit, one of the three toolkits based on OPEN LOOK interface specification defined by SUN and AT&T. This toolkit is derived from the previous SunView interface, but it runs on X. They have this toolkit in mind for customers who want to migrate from SunView toolkit.

AT&T meanwhile offers their OPEN LOOK Toolkit for beta test now. This toolkit is called Xt+.

SUN also offers a third OPEN LOOK Toolkit based on their Postscript Window System NeWS on X.

2. There was an interesting talk by Silicon Graphics on the problem they have in trying to implement X server on their specialized hardware: a workstation based on pipelined display architecture. It is an art of compromising. They try to follow X design philosophy and at the same time try to optimize X server based on special hardware. Most of these issues are obviously old, *i.e.*, known by the X WINDOWS Consortium, but there is no standard approach suggested to implementors yet.

3. Hewlett Packard has a new laboratory called the Multi-User Laboratory to explore computer capabilities in a supported environment in which a group of people communicate with each other, coordinate their activities, and share information. They are using X WINDOWS and their "video-network" *to transmit multi-media information.*

4. It is really a vote of confidence for the X WINDOW System when it was shown that X server has been implemented in some hostile languages (from its standard implementation in C). There were two talks about the experiences in this implementation work: one using the Cedar language, the other based on Ada.

5. A person from Stanford University also presented his development environment called ADEW based on the ANDREW toolkit (which in turn is implemented on X) of Carnegie Mellon University. One thing I observed in this conference is that there is no lack of sophisticated toolkits for application development on X. The terminology ranges from *toolkit*, to *widget*, to *development environment*. I am aware of at least of a dozen of such development tools from both industry and universities. Each of them tries to promote a different "look and feel" and different development methodology. The great things about these tools are: almost all of them are free with sources, they make the application development work much easier by providing a high level abstraction to user-interface implementation, and they efficiently package the layer of Xlib calls to reduce communication traffic.

6. Finally, I attended a Bird-of-Feather session in the evening on the topic of Server Implementation. Bob Scheifler (the MIT X Consortium director) was there to answer questions. I had good impressions about how developers feel and the consortium plans for the future. They agreed to work on some critical issues of packaging, documenting the source codes to make the installation, and modification easier for server implementors. In response to a remark that X appears to be slow, Bob said that it is not true in general, it is just unfortunate that the implementation on the SUN workstation was [in that particular case] not efficient and that the SUN is the most popular workstation.

IV.

ADDENDUM

The following concerning commercial application developments was taken verbatim off the USENET, and

is included here for reference. This is used as an example of developments taking place during 1988 as a result of the commercialization of the MIT X WINDOWS development.

Subject: Fw: NCD16 X Window Display Station
Article 185 of comp.newprod:
Path: daitc!vrDXhq!grebyn!haven!rutgers!gatech!mcdchg!usenet
From: mikeh@lupine.UUCP (Mike Harrigan)
Keywords: X Window System Workstations Displays
Date: 25 Jan 89 17:30:43 GMT
Organization: Network Computing Devices, Palo Alto, CA

X WINDOW DISPLAY STATION FROM NETWORK COMPUTING DEVICES COMBINES WORKSTATION-STYLE GRAPHICAL USER INTERFACE WITH TERMINAL ERGONOMICS AND AFFORDABLE PRICING

MOUNTAIN VIEW, Calif., Jan. 23, 1989 -- A new type of network peripheral that provides workstation-like display capability -- including a windowed interface, high resolution and advanced communications -- for about \$2,500, or half the entry price of a typical workstation, has been introduced by Network Computing Devices, Inc.

The NCD16 Network Display Station, the year-old firm's first product, is a compact desktop unit based on the X Window System, the de facto industry standard for high-performance windowing applications. It offers a sophisticated yet affordable user interface to networks of computers supporting X, including those running the UNIX and VMS operating systems.

Running the industry-standard TCP/IP communication protocols, the NCD16 is a general-purpose product suited for a wide variety of engineering, scientific and professional applications, such as software development; computer-aided design, manufacturing and engineering; information retrieval; transaction processing; and computer-aided publishing.

Filling the Terminal-Workstation Gap

The NCD16 "defines a new category of product designed specifically for network computing environments, where an ASCII terminal doesn't do enough and a workstation is overkill," said Judy Estrin, NCD executive vice president.

"Network users, particularly in the engineering and scientific community, typically run complex applications on UNIX- or VMS-based hosts and need a sophisticated display function at their desks to output the results of those applications," Estrin said. "To get a display with multiple windows and high resolution, these users have had to buy PCs -- and add high-resolution monitors -- or diskless workstations at \$5,000 and up. But a large part of what they're paying for is the machine's application processing power, which often goes unused."

Dedicated to display and communications functions rather than application, the NCD16 fills this gap between the low-function terminal and the high-priced workstation, she said. Its workstation-style user interface can access multiple hosts and maintain multiple active sessions, communicating the data for each session to the user through a separate display window.

A Graphics Front End for Minicomputers

"An ideal application for the NCD16," Estrin said, "is as a graphics front end to VAXes or other minicomputers, which have suffered from lack of the windowing capability usually associated with workstations. The NCD16 lets users protect and leverage large installed bases of minis without making costly investments in workstations and PCs."

The NCD16 Network Display Station includes a monochrome monitor, single-board electronics located in the monitor's base, keyboard, mouse, software and two communications interfaces.

The unit's unique 16-inch diagonal, square-format monitor offers 1024- by 1024-pixel, 105-dot-per-inch resolution. "You get the same million-pixel functionality as on the 19-inch monitors typically used with workstations," Estrin said. "But with a footprint of only 13 by 13 inches for the monitor and base, the NCD16 take up only about half the space on the desktop."

Incorporating the latest ergonomic standards in monitor technology, the NCD16 has a fast 70-Hz refresh rate for flicker-free operation, and supports and "overscan" features which eliminates the distracting black border that surrounds the image on most displays.

Quiet Fanless Operation

Because the NCD16 is a compact product dedicated to display functions, without the PC or workstation's requirements for internal expansion space, it can use a convection cooling system, eliminating the need for a fan and ensuring noiseless operation.

The NCD16 is based on a 12.5-MHz MC68000 microprocessor which runs the X Window System software and the TCP/IP protocols. A graphics co-processor assists in display functions. The board also contains 1 to 4.5 megabytes of dynamic random-access memory (DRAM).

Modular Network Interface for Future Upgrades

The NCD16 comes with one asynchronous RS232 serial interface, and offers the user a choice of an Ethernet (IEEE 802.3) and thin Ethernet interface, or alternatively a second RS232/RS422 port. Both interfaces are contained on a removable module that plugs into the NCD16 processor board and can be exchanged later as interfaces to other network technologies (e.g., ISDN, FDDI, Token Ring) are made available. The RS232 port support SLIP communications at speeds up to 38.4 Kbps.

The NCD16 supports both the TCP/IP Domain and IEN 116 Name Servers, which permit access to networked devices by name rather than numerical address.

The use of standard input device interfaces -- a PS/2 interface for the NCD16 keyboard and a serial interface for the mouse -- ensures that the user can easily substitute preferred devices if desired.

X Window Server Software

The X Window System standard, developed by the Massachusetts Institute of Technology and supported by more than 35 major computer vendors, permits the division of an application program into two parts: user interaction (or display server) and computation. Dedicated solely to display functions, the NCD16 supports the display server portion of the software, which enables it to provide display services for any networked host that supports X Window System applications. NCD has optimized the software for performance on the NCD platform.

Software is downloaded into the NCD16 from any networked host that supports the TFTP file-transfer protocol, facilitating fast and efficient software updating; a PROM-loading option also is available. A configuration and setup utility is included with the unit.

Pricing/Availability

Available 60 days after receipt of order, the NCD16 is priced starting at \$2,550 with monitor, base, keyboard, mouse, software, one RS232 serial port and one Ethernet interface.

Estrin said the NCD16 is the first member of a family of network display stations. Later versions will feature different screen sizes, color graphics and increased performance.

Network Computing Devices was formed in February 1988 by three former executives from Ridge Computers; the founding team was joined in July by William Carrico, president, and Judith Estrin, executive vice president, co-founders of Bridge Communications and later executives of 3Com Corporation. The company has raised \$5 million in startup financing from four venture capital firms.

Company contact:

Judy Estrin
Network Computing Devices
(415) 694-0650 or lupine!info@uunet.uu.net

Hypermedia Laboratory technology watch:

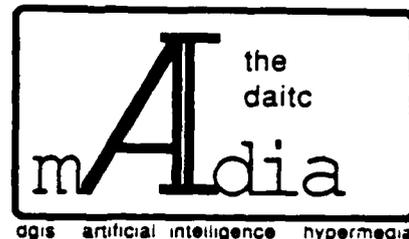
Date: Tue Jan 3 11:28:01 1989 From: duc (Duc Tran) Subject: X on Mac

(forwarding comment)

The import of the following would be having a DGIS >>standard<< hypermedia-based (bitmap systems, images, processes) running off the DGIS for >>diverse<< bitmap systems.

(original message follows)

I thought this new product is interesting: it's called **Exodus**. It runs under Macintosh **Multi-Finder** and it will turn any MAC into an X WINDOW terminal. Other MAC windows can coexist with this software running. With this software we can have a UNIX box, like our Pyramid super-mini, tell X WINDOW terminals, such as SUN, MAC, Symbolics, ... to all display in the same way. With Exodus an application programmer now need write applications with X user-interface only once. He then can select to use the applications with any X WINDOWS terminal as implemented on these popular boxes without having to modify his code! It costs in the area of \$500.



THE SYMBOLICS PROCESSOR

The Symbolics 3650 Processor in the DAITC Hypermedia Lab.

The Symbolics is a high-end list processing (LISP) workstation that has advanced development tools at its disposal. The power of the Symbolics is its Genera Software environment. This environment is written completely in LISP, (except for a small amount of machine micro-code), which makes writing symbolic oriented applications very natural and efficient. UNIX-based machines restrict AI development in terms of how the particular language handles interfacing with the UNIX internals; the LISP machine is open to direct manipulation from the LISP programmer. This is a highly desired environment to the developer. Adding to the inherent power of list manipulation on the Symbolics is the library of development tools residing on our workstation, which includes the following:

- * KEE - Knowledge Engineering Environment, by Intellicorp, is an Expert System shell tool which gives us the power to rapidly engineer an Expert System for any specific topic of concern. The knowledge engineer is tasked with collecting the knowledge, consisting of rules and data, and then inputting it into the framework of KEE's rule and knowledge system. KEE lets you incorporate graphics as well as text to enhance the user interface with the Expert System. A rapidly developing concept for the Hypermedia Lab use of KEE would be to create a DAITC Network Expert System for use by systems staff as well as DAITC management to analyze and manage the DAITC network.

- * CLOE/ PC386 Environment - With the presence of a 386 processor and the CLOE software in the Symbolics we are capable of developing applications on the Symbolics and porting them to a regular 386 based personal computer running DOS or UNIX. The ability to develop UNIX and DOS based applications opens an endless amount of doors on PC based Artificial Intelligence applications. We can capitalize on the power of the Symbolics to create applications for the IBM PC computer line. The strong debugging environment on the Symbolics is one of the major strengths in this method of development, since PCs have very few powerful debugging tools. PCs are the way of the future and we have a way with CLOE/386 to quickly and powerfully prototype ideas for these machines.

- * STATICE - Stalice, by Symbolics, is an Object Oriented Database Management System (OODBMS) for the Symbolics environment. OODBMS's and Object Oriented Programming are very powerful newly recognized concepts in engineering and programming circles. The idea is to design systems around "objects" and not around data types or data tuples. By looking at problems as a collection of related objects one approaches more closely the actual problem solving techniques of the human mind. These OODBMS and OOPL's provide the means to follow an idea through from conception to implementation using object oriented techniques. Therefore, with STATICE we will be able to design and implement an object oriented database and combine this with LISP object oriented programming using FLAVORS. FLAVORS is the object oriented portion of the GENERA environment enabling the developer to treat LISP code as objects.

- * DOCUMENT EXAMINER - The DOCUMENT EXAMINER, by Symbolics, is a frequently cited hypertext application which provides for the user online access to the 10 Symbolics manuals in a non-linear fashion. This serves as a powerful example for future hypermedia developments as well as providing an improved state of development from rapid manual access.

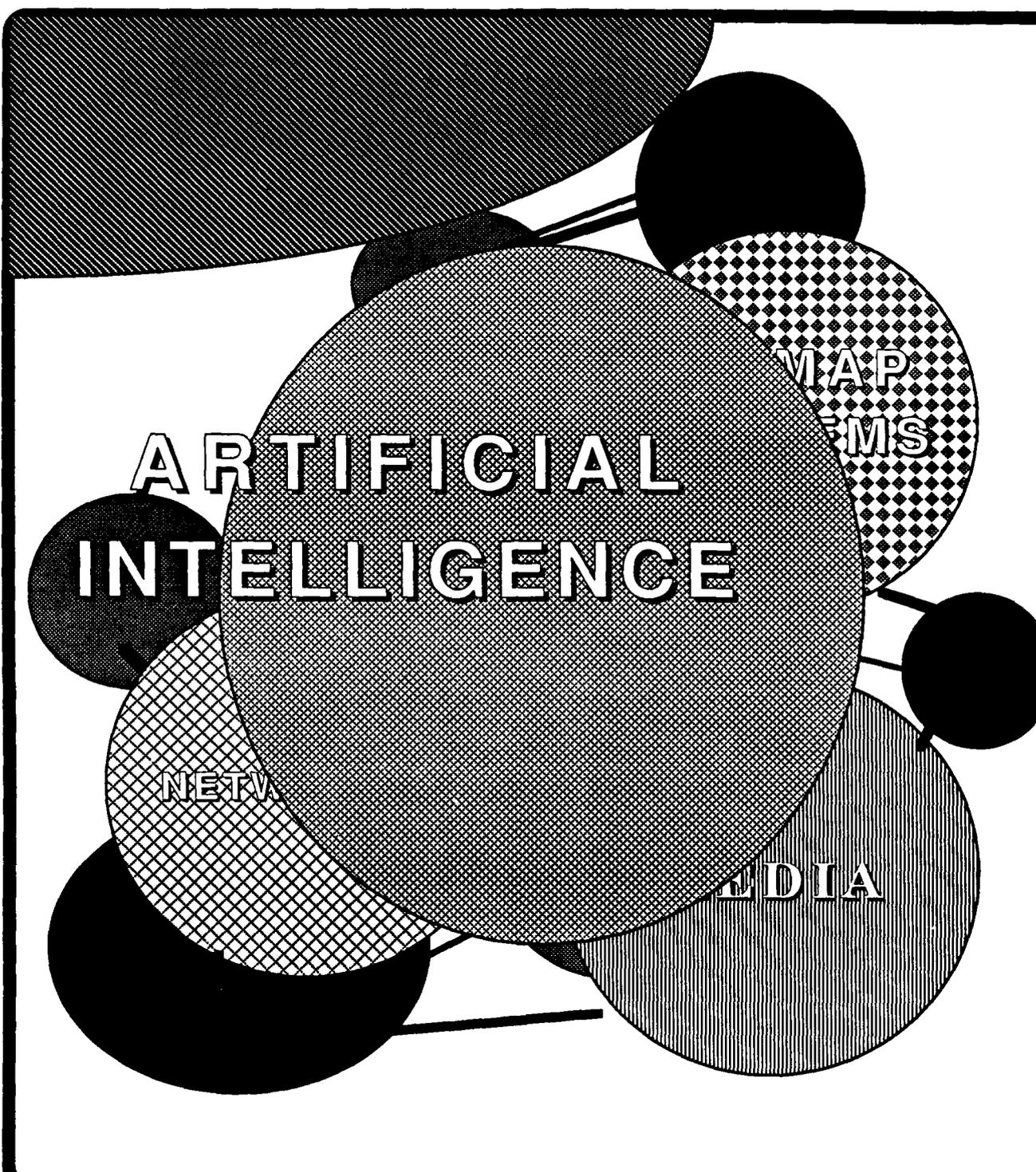
- * Debugging Environment - The interactive debugging environment on the Symbolics provides for a very powerful and informative development testing phase. This decreases development time needed to take an idea to the prototype stage.

- * PROLOG and C - We are capable of developing applications utilizing the strengths of languages like PROLOG and C as well as LISP. Again rapid prototyping of systems written in these languages is possible with the strength of the Symbolics debugging tools.

There are many less dramatic characteristics which make the Symbolics such a powerful development machine besides the previously mentioned items. But all lead to the conclusion that the Symbolics is a very useful machine in development arenas. One of the DAITC's main goals is to develop prototypes to help find solutions to the DOD's information management requirements. This prototyping is best accomplished, economically and efficiently, by using the many tools of the Symbolics environment.

The prime goal of the Hypermedia Lab. is to create AI-based programs that make the human-machine interface more human in the machine's responses.

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER (DAITC), ALEXANDRIA, VIRGINIA

An abstract graphic design featuring several overlapping circles and a large rectangular shape at the top left. The circles are filled with different patterns: a solid black circle, a circle with a fine grid pattern, a circle with a larger grid pattern, and a circle with vertical lines. The text 'ARTIFICIAL INTELLIGENCE' is centered in a large, bold, sans-serif font. Other text elements include 'MAP' and 'EMS' on the top-right circle, 'NETV' on the bottom-left circle, and 'EDIA' on the bottom-right circle. The entire design is enclosed in a thick black border.

**ARTIFICIAL
INTELLIGENCE**

MAP
EMS

NETV

EDIA

DAITC HYPERMEDIA LABORATORY

**TOWARD AN ARTIFICIAL INTELLIGENCE ENVIRONMENT FOR
DEFENSE TECHNICAL INFORMATION CENTER (DTIC):
PROPOSED TASKS -- 1987**

*Allan D. Kuhn
Defense Technical Information Center*

The following is extracted from a DTIC document (1) issued in May 1987. That document presented an outline of potential artificial intelligence (AI) development activities relative to the DoD Gateway Information System (DGIS). This outline thus basically laid out a plan for initiating AI implementation by DTIC.

I. INTRODUCTION

A major element to be considered for incorporating AI tools is the configuration of our base system, the DoD Gateway Information System (DGIS). The DGIS, which is the core computer system of the Defense Applied Information Technology Center (DAITC), serves two basic purposes. The first, the mission of the DGIS, is as a service to the DoD community for accessing the plethora of information systems, aggregating information from them, and processing the aggregated information into a product that is useful to the user.

The second purpose of DGIS is as a prototype environment for enhancing information access for the DoD community. The uses and configuration of the system are mutually influential, with needs determining configuration, but configuration determining applications.

This paper explores AI development tasks relevant to DGIS.

II. PROPOSED TASKS

This section describes effort which have been identified as candidates for application of AI technology. The main criterion for identifying such efforts is the beneficial effect on and for the user. The application of both AI and AI-like technology has the effect of making the electronic system appear to have human-like responses. These "human-like" responses make it easier for DoD managers to successfully interact with and extract relevant data from information systems. DTIC's first excursion into the AI arena involved the acquisition of PROLOG for the DGIS Common Command Language (CCL) project.

DTIC has experienced first-hand the dendritic effect of venturing into AI: requirements and applications will generate as developments progress. For example, in the case of the CCL project, the incorporation of PROLOG immediately allows the the creation of a command language knowledge base system (KBS) as a searcher assistant. KBS development, however, leads to the need for an expert system (ES) building tool; such a tool not only represents higher level AI application, but also makes use of a developed application that is recognized as current technology. The inclusion of KBS/ES technologies generate further highly useful applications, which in the example of CCL is a CCL user profile KBS. This KBS will permit the user to tailor a CCL query system to meet one's particular needs. Not the least is the eventual employment of natural language in eliciting and formulating the user's query.

The very basic goal in implementing AI applications, then, is to make the human-machine interface more human-like, as indicated by the following tasks.

1. COMMON COMMAND LANGUAGE

There are two primary elements that establish the need for an AI-based common command language (CCL) on the DGIS. The first is to be able to "talk" with a multitude of databases without needing to know the different native command languages for them. The second is being able to address databases without needing to know their individual search [behaviors]. An algorithmic program, e.g., C, can be written to create a CCL capability as a simple substitute for native command languages. The varying [retrieval requirements of the different systems], however, have generated the potential of incorporating AI capabilities. An AI-based CCL, working out of knowledge base systems for command languages and user profiles, can create a human-machine interface that is more human in its processes.

2. DIVERSE DATABASE QUERY EXPERT SYSTEM

The refinement of the query ("*Is this what I am really asking?*"), and transmission of the query to invoke the search relative to the database functionality. This system would include the following subsystems:

a. Query Analysis System -

System for refining the user's search query, patterned after standard reference librarian query refinement models, in natural language.

b. Subject Searching System -

Extension of the Query Analysis System, making use of multiple database thesauri.

c. Full Text Search System -

Inclusion of text search applications as a component; it would be used for retrieving not only full text stored documents, but also on all text data, e.g., the titles and abstracts of bibliographic databases. Searching documents in their language implies the need for making use of the the user's language, i.e., natural language querying; this in turn permits greater specificity, which results with increased topic relevance and context. The system needs to provide for retrieving user-designated size sections containing the user's topic in context, for browsing documents online, and capabilities for analysing and processing text information.

d. Information Scan System -

The scanning of user-selected sets of databases for query results, to determine the results volume and consequently the usefulness of each database in the set pertinent to the query.

e. Trending and Projecting System -

Statistical-based trending, along the lines of management trending and projecting packages, of user-selected databases, including in sets. For example, comparison trending and projecting would be made on two or more terms or topics.

3. THESAURI INTEGRATION FOR EXPERT SEARCHING

Development of a system to apply external database thesauri for determining relevant databases in response to a user's query.

4. INFORMATION PROCESSOR SYSTEMS

After information has been aggregated from diverse sources, the user needs the capability to standardize, process and reformat that aggregation into a product that is most useful to him. DGIS already has these capabilities to a certain extent, programmed in C. AI-based systems that would enhance and expand the

quality of DGIS information processing would include a duplicate and irrelevant data analysis system, a relevant data analysis system, analysis graphics, and a tailored product system. These systems would comprise the whole of user-invoked DGIS information processing as follows:

a. Translator Generator -

An expert system for both programmer and user creation of a generator for translating an external database record format hitherto not existent in DGIS, into the DGIS standard format. Initially, this generator might be constituted of a selection of primary fields, based on the information displayed in the final product.

b. Duplicate and Irrelevant Data Analysis System -

Realizing that aggregation of information on a selected subject from diverse but pertinent information systems can result in redundant citations, this system would not only recognize redundancies, but also allow the user to determine which citations take precedence for remaining in the aggregation. This system should also recognize the lesser occurrence of irrelevant citations based on subject disjuncture with the searcher's original query.

c. Relevant Data Analysis System -

The system for doing analysis routines on the aggregated record query relevances, invoked at the discretion of the user. These routines, based on those currently available in DGIS, would perform statistical and text analyses that are bibliometrically useful to the user. AI-based routines would expand and deepen the functions of current C-based routines.

d. Information Analysis Graphics System -

In that a picture to the human mind is worth a thousand words, analytical graphics capabilities should be incorporated. ASCII-based graphics representation should also be included for those users not having bit-map terminals. Examples of statistical graphics are: occurrence of records by year, author, subject term; graphics comparison of occurrences; bar, pie, and line charts; et al.

e. Tailored Product Generation System -

This system would tailor presentable information products, in formats ascertained by the enduser. Such a system might be considered a desk-top publishing system for 1-n copies of a document, to include the standard components of a document such as cover, content, tables, graphics, and indices.

5. ROUTINE GENERATOR EXPERT SYSTEM

This is an expert system for creating generators for routine processes. This system would be used by both programmers and users, as appropriate. The expert system must make use of bidirectional natural language to define the routine, state the requirements, and build the generator. Examples of identified routines for which generators might be created are:

o Connection Agent Generator - To establish connects to remote databases not yet in DGIS, but needed by a particular user on a recurring basis. If the connect would not be incorporated in the DGIS connection routines, it could at least be retained by the user.

o Translator Generator for Common Command Language - To create the translation of a remote database command language that has not yet been done for DGIS, but is needed by a particular user. A user-generated translation could initially be based on a set of basic commands most commonly used. The ability to toggle back and forth between CCL and the native command language would be automatically included.

o *Translator Generator for Information Processing* - Discussed above in section on Information Processing Systems.

6. FOREIGN HUMAN LANGUAGE INTERFACES

The incorporation of human language programs that provide passive and interactive machine translation and transmission, as follows:

- a. Translation of foreign language information retrieved from foreign databases.
- b. In-process translation into English of messages received on the electronic mail from foreign language speaking people, with reciprocal translation from English.
- c. Communicating interactively with foreign language databases, with interactive translation between English and the language of the database or record.
- d. Interactive translated communication with foreign language speaking people while making use of the DGIS capability to "talk" interactively via the keyboard. An extension would be translation of voice-generated data.

7. NUMERIC INFORMATION QUERY AND PROCESSING SYSTEM

A system for identifying numeric information sources, and aggregating, analyzing, and synthesizing that information.

8. INFORMATION ANALYSIS AND EXPERT SYSTEM APPLICATION SYSTEM

A system that analyzes the information content of a targetted universe, clarifies the purpose of the intended Expert System, and recommends the standard categorical Expert System functionality. This system would be used in the building of expert systems as needed for the DGIS, plus expert system developments made for application outside DGIS.

9. PORTABLE NATURAL LANGUAGE INTERFACE

Natural language interface commonly usable with programs, routines, functions, and other areas and aspects of the electronic system.

10. NATURAL ENGLISH LANGUAGE INTERFACE TO UNIX

Natural English language interface to cryptic UNIX, to include UNIX function identification by the user's natural language description.

III. RECOMMENDED CONFIGURATIONS

The DTIC AI environment is by no means to be considered an AI research activity but an AI applications activity. Our purpose is to take both developed and leading edge applications for enhancing information accessing for our users. The following discusses configurations for mission, organization, and system.

1. MISSION CONFIGURATION

It is recommended that the DTIC AI activity mission be:

- a. Identify DoD information problem areas,

- b. Evaluate those problem areas for potential AI applications,
- c. Build applications using currently available ware while at the same time using leading edge technology,
- d. Implement those applications in a timely manner.

The DTIC AI applications development will have a product orientation. The quality and reliability of the activity will be based on its products. It is recommended, therefore, that the first activity be to construct a relatively easy-to-build system as rapidly as possible for the purpose of not only having a product, but also a demonstrable system that shows the usefulness of AI applications. This initial system would also serve as a learning exercise in product building.

The AI activity should be staffed based on a set of high quality criteria in order to excel in its mission. The activity should also form liaisons with other DoD and government AI activities, for awareness and cooperation on AI development and implementation, especially in information.

2. ORGANIZATION CONFIGURATIONS

DTIC has two options currently [1987] available for an organization configuration:

- a. DTIC-sponsored and monitored AI development activity.

DTIC has already begun AI activity on DGIS through the Common Command Language effort. This development opened the door to AI for DTIC, beginning with the acquisition of PROLOG for installation on the DGIS. The potential for expanding AI development is limited primarily by funding constraints.

- b. DTIC participant in the DAITC AI laboratory.

The Defense Applied Information Technology Center (DAITC), of which DTIC is a sponsor, and for which DTIC is to provide the Federal Government management, has already established an embryonic AI Laboratory. Contributing to this resource and taking full advantage of it would be both economical and cost effective. It is planned that eventually DTIC will provide the Federal manager for this Laboratory. Once that is achieved, the AI activities of the DAITC Laboratory and of DGIS could be merged and coordinated to improve cost efficiencies and accelerate progress.

The cost-benefit of a joint DTIC-DAITC AI laboratory, with the sharing of its resources, makes this option an obvious recommendation.

=====

V. SUMMARY

1. PLAN AND MISSION

A DTIC goal is to increase database services directly to endusers. DTIC's endusers are the research and development (R&D) managers, engineers, and scientists. These users depend on accurate and up-to-date information to fulfill their mission assignments. Proliferation of microcomputers makes database access feasible for the DoD enduser.

A major obstacle for enduser database utilization is knowing what and how to search. These users need a system that will allow them to identify, access, and query databases using natural language dialogues. Artificial intelligence technology applied to information retrieval systems provides potentials to fulfill these user requirements. The establishment of an Artificial Intelligence environment will be the mechanism for

the development and incorporation of AI applications in the DoD Gateway Information System and wherever else needed.

The AI activity needs both expertise and materiel to function. The AI expertise required should at the minimum be at a Master's level with extensive C/UNIX and AI programming backgrounds. A PhD. level expert is to have this level of studies in AI applications relevant to the needs of the DGIS programs. The hardware and software requirements . . . will allow development of programs incorporating [for example] the use of expert systems and use of natural language as a means of providing the needed information querying capabilities to endusers. AI activities, after they have developed, could further lead into exploring other DoD enduser needs such as technical information networking and transfer, intelligence programs, and industrial/manufacturing programs assessment.

2. BENEFITS

The application of AI to information retrieval will allow DoD R&D managers, engineers, and scientists to overcome the barriers of user-system communications. They will be able to more easily access and utilize the information they need. The availability of this information will make our endusers more efficient and help them avoid the high costs of duplicating or ignoring pertinent work.

The incorporation of AI applications in DGIS will additionally make the human-machine interface on DGIS more human-like in its functions and responses, and tolerant of human frailties that are caused by the complexities of the human mind.

=====

EXTRACT REFERENCE

[1] *Toward an Artificial Intelligence Environment for DTIC: Proposed Tasks, Recommended Configurations, Projected Start-Up Costs.*
Allan D. Kuhn, Defense Technical Information Center, May 1987, AD-A181 103.

BACKGROUND REFERENCES

Artificial Intelligence Developments Re: DoD Gateway Information System (DGIS) & Defense Applied Information Technology Center (DAITC).
Kuhn, Allan D. Defense Technical Information Center, February 1987, AD-A181 101.

The DoD Gateway Information System.
Cotter, Gladys A. Defense Technical Information Center, October 1985, AD-A161 701.

The DoD Gateway Information System: Prototype Experience.
Cotter, Gladys A. Defense Technical Information Center, April 1986, AD-A166 200.

Toward an Artificial Intelligence Environment for DTIC: Staffing Qualification Criteria for AI Application Development.
Kuhn, Allan D. & Tran, Duc T. Defense Technical Information Center, February 1987, AD-181 100.

**THE DGIS COMMON COMMAND LANGUAGE SYSTEM (CCLS)
DGIS STEPS INTO ARTIFICIAL INTELLIGENCE**

*Allan Kuhn
Hypermedia Laboratory*

The section following is extracted and reproduced from:

Kuhn, Allan D., and Gladys A. Cotter. "The DoD Gateway Information System (DGIS): The Department of Defense Microcomputer User's Gateway to the World;" *Microcomputers for Information Management: An International Journal for Library and Information Services*, v5 n2, June 1988, pp. 73-92. This article is available from DTIC or NTIS as AD-A203 351.

Background references for this extract are:

Bixby, Randy L. *DoD Gateway Information System (DGIS): Common Command Language Mapping*. DTIC, October 1987, AD-A185 951.

Kuhn, Allan D., Randy L. Bixby, and Duc T. Tran. *DoD Gateway Information System (DGIS) Common Command Language: The First Prototyping and the Decision for Artificial Intelligence*. DTIC, August 1987, AD-A185 950. Presented at and published *sans addenda* in Proceedings of the Ninth National Online Meeting, New York, May 10-12 1988.

Kuhn, Allan D., Randy L. Bixby, and Duc T. Tran. *DoD Gateway Information System (DGIS) Common Command Language: The Decision for Artificial Intelligence*. DTIC, March 1988, AD-A197 215. Presented at and published in RIAO88 Conference on User-Oriented Computer-Based Text and Image Handling, Massachusetts Institute of Technology, Cambridge, March 21-24 1988.

National Information Standards Organization (NISO). *American National Standard - For Information Sciences: Common Command Language for Online Interactive Information Retrieval (DRAFT)*. March 1987, ANSI Z39.58-198X.

Tran, Duc T. *DoD Gateway Information System (DGIS) Common Command Language: PROLOG Knowledge Base Profile*. DTIC, October 1987, AD-A186 150.

DGIS STEPS INTO ARTIFICIAL INTELLIGENCE

A. Common Command Language

Our entry into AI began with our development of a common command language (CCL). DGIS is creating a CCL system to serve the skillful searcher access multiple information systems, including those with which the searcher may not be familiar. Our CCL system is based on the draft standard put out by the National Information Standards Organization (NISO, 1987).

Originally we thought CCL might be accomplished with traditional third generation programming (in this case, C language). We developed our first prototypes with DIALOG, BRS, NASA-RECON, and DROLS (Defense RDT&E On-Line System). These first prototypes made us realize that information systems as a whole represent a very heterogeneous universe. The C language prototypes, therefore, served as a study stage (Kuhn, Bixby, & Tran, 1987; Bixby, 1987) for the problems and feasibility of CCL, during which we also began looking at AI as a means to handle the complexities of information systems and their operational behaviors.

We selected PROLOG as the appropriate programming language to incorporate AI capabilities into CCL and also initiate AI-based programs into DGIS. PROLOG was chosen because of its logic programming characteristics and tools

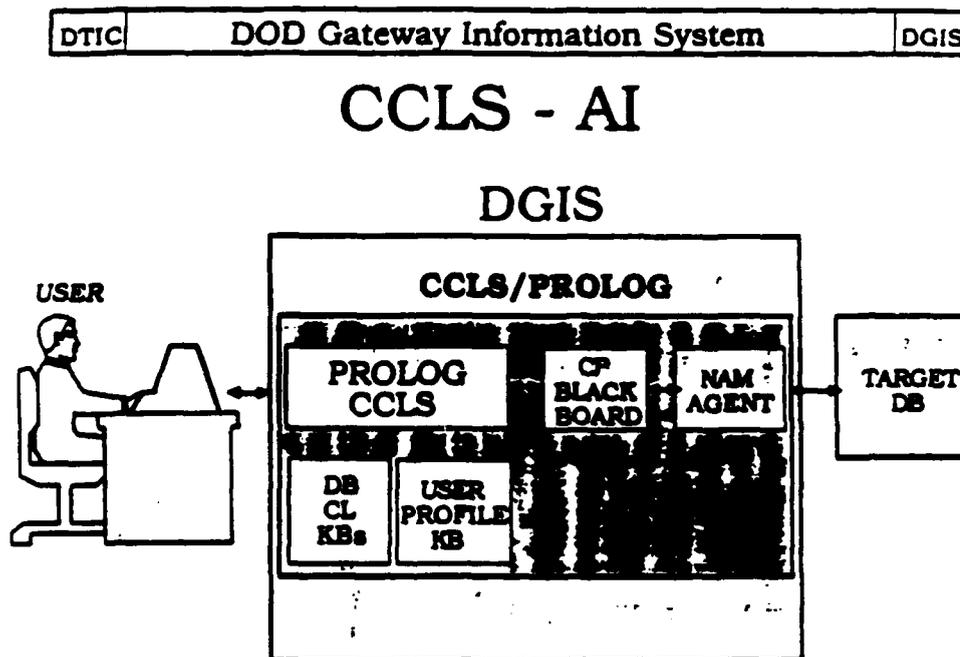


Figure 2. The Architecture of DGIS CCLS-AI.

```
DGIS/vax(1)% ccl

*** Welcome to DGIS CCL ***

This system is the first AI-based component
on DGIS to give the user standard access to
diverse information systems.

The following systems are currently available:
dialog.brs.drots.nam

Please specify a CCL system ('D to exit) > dialog

Trying to establish connection to DIALOG using NAM.

Attempting hardwired connection at 1200 baud to TYMNET. Connection established to TYMNET.
Attempting TYMNET connection to DIALOG2. TYMNET connection established to DIALOG2.
Attempting login. Login complete.

*****
Loading DIALOG knowledge base ...

Loading User Profile Knowledge Base ... [/ai/ccl/proto/uxsth/uxtha.pl consulted (0.300 sec 340 bytes)]

CCL > explain

The EXPLAIN command provides information on various topics in CCL DIALOG. The following topics
can be explained:

choose display find show stop
define explain general start

CCL > etc.
```

Figure 3. DGIS CCL prototype: Initiating DGIS CCL. Note banner, systems available, connecting to target system in CCL mode, login, DIALOG Knowledge Base loaded, User Profile Knowledge Base loaded, entry of NISO common command "explain" in place of "help."

such as its inherent reversibility, database capabilities, separation of logic and control, and its object inheritance and message passing. PROLOG additionally fits well with the intended plan for a gradual migration of CCL from a structured command language to natural language. The current phase of the project utilizes blackboard architecture and knowledge-based driven knowledge sources. Figure 2 illustrates the architecture of the CCL system. The "NAM agent" in the figure is the connection agent; NAM stands for Network Access Machine, which is in

reality connection agent software developed originally by the National Bureau of Standards.

Figures 3 through 7 show the actual screen dumps on how the user can sign on to various information systems through DGIS CCL. Figure 3 illustrates how DGIS CCL is initiated and how DIALOG knowledge base is accessed. Figures 4 and 5 show how "EXPLAIN," "CHOOSE," and "SHOW news" work in DOD-DROLS. Figure 6 illustrates how a sample CCL search and retrieval session works in NASA knowledge base, and finally, Figure 7 provides a sample display of how a CCLS session is opened and closed.

DGIS CCL, therefore, is being structured as a knowledge-based system. The control program is a blackboard-based architecture PROLOG program that controls the interaction between the CCL agents and the communication agents, or knowledge sources. The CCL knowledge sources are the experts based on a number of knowledge-based systems. Typically, there are two types of CCL knowledge bases: One is pertinent to user information, and the other is knowledge about databases. The user knowledge base system stores information relevant to a particular user or a group of users. The database knowledge base contains information needed to translate CCL commands into host database

```

Loading DROLS knowledge base ...
-----
CCL > explain

The EXPLAIN command provides information on various topics in CCL DROLS. The following topics
can be explained:

choose define display explain find show start stop

CCL > explain choose

The CHOOSE command identifies the database to be searched on DROLS. The format of the CHOOSE
command is:

    CHOOSE <database-id>

Three files are currently available on CCL DROLS. These are:

    Technical Reports File (TR)
    Work Units Information System (WU)
    Current Reports File (CF)

DROLS has no default database. CCL, however, has defined the Technical Reports (TR) File as the default
file if the CHOOSE command has not been invoked. If you would prefer to change that default to another
file more convenient to you, you may do so in the User Profile Knowledge Base.

CCL > choose -ez.-

```

Figure 4. DGIS CCLS prototype: The "explain" for "choose" in DoD-DROLS; note reference to including a user-chosen default database in the User Profile Knowledge Base.

```
Loading DROLS knowledge base ...
-----

CCL > show

The following topics are available for SHOWing:
news time cost find choose display

CCL > show news

Currently there four CCL prototype systems :

      BRS, DIALOG2, DROLS, and NASA

The following commands are possible in these prototypes

      CHOOSE FIND DISPLAY

and the support commands:

      EXPLAIN DEFINE DELETE SHOW START STOP

Comments and feedbacks are sollicitized and please send them to

      bixby@dgis and kuhn@dgis

CCL > -etc.-
```

Figure 5. DGIS CCLS prototype: "show news" for DGIS CCLS.

commands and to understand the returning results and errors from the database. The control program is an object-oriented blackboard-based program in which the blackboard is nothing more than a general object that registers and monitors progress of the related knowledge sources (Tran, 1987).

The very basic problem of CCL surfaced as our experience grew. We learned that creating "Common Command Language" was *not* a panacea; programming a "standard" command language was in actuality only substituting one command language for another. The idiosyncracies of information systems include individualistic formats for displaying information, distinctive options for information selections, and a lack of functional information that an expert searcher knows but is not indicated on the system. The major effort involves establishing a CCL system that aids the searcher in searching unfamiliar information systems as well as the familiar ones. Our implementation of AI-based tools to do this will hopefully achieve this universality.

```

DGIS/vax(1)% ccl
-----
Loading NASA knowledge base ...
Loading User Profile Knowledge Base ...
-----
CCL > choose
NASA-RECON identifies its databases by number or by letter. Please consult system documentation for
proper database identification.
Illegal CCL command !!
CCL > choose a
You have chosen NASA-RECON file a
CCL > find
select [] <<<<This is the echo target database (NASA) entry
SELECT
TERM IN SELECT COMMAND NOT IN DICTIONARY
ENTER:
CCL > find artificial intelligence
select artificial intelligence <<<<echo
1 2718 2718 ST/ARTIFICIAL INTELLIGENCE
ENTER:
CCL > display
Invalid CCL command.
CCL > explain display
The DISPLAY command initiates an online display of results from the immediately preceding search
results set. Three DISPLAY commands are currently available on CCL NASA-RECON. These are: -etc.-
CCL > display short 1-3
The DISPLAY command in CCL NASA requires specification of a set number. Please re-enter your
DISPLAY command in the format:
DISPLAY <set no.> <format> <term>
CCL > display 1 short 1-3
browse 1/6/1-3 <<<<echo
BROWSE 1/6/1
88A29425 ISSUE 11 PAGE 1738 CATEGORY 61 87/00/00 176 PAGES in
RUSSIAN UNCLASSIFIED DOCUMENT COPYRIGHT
UTTL: Pattern recognition and natural language understanding by a computer --
Russian book
AUTH: A/PAIN, VITALII SAMOLOVICH
-etc.-
-----
CCL > start
Current system must be terminated before another system can be started.
CCL > stop
*** Goodbye ***
[ End of Prolog execution ]

*** GOODBYE ***

DGIS/vax(2)%

```

Figure 6. DGIS CCL prototype: a sample CCL search and retrieval session, including mistaken entries with CCL responses.

```
DGIS/vax[1]% ccl
*** Welcome to DGIS CCL ***
-----
Please specify a CCL system (^D to exit) > brs
-----
Login Completed
Loading BRS knowledge base ...
Loading User Profile Knowledge Base ...
[/ai/ccl/proto/userkb/kuhn.pl consulted (0.300 sec 340 bytes)]
-----
CCL > -CCL session-
CCL > stop
*** Goodbye ***

[ End of Prolog execution ]
*** GOODBYE ***
DGIS/vax[2]%
```

Figure 7. DGIS CCLS prototype: example display of opening and closing a CCLS session.

ENHANCEMENTS OUTLINE FOR THE DOD GATEWAY INFORMATION SYSTEM (DGIS): PROPOSAL FOR AI/ES-BASED INFORMATION MODELLING TECHNOLOGY

Richard S. Marcus
Lab. for Information and Decision Systems, Massachusetts Institute of Technology

April 1988

A. BACKGROUND

The Marcus report *Experimental Evaluation of CONIT in DGIS Gateway Environment* (February, 1988)* concluded that techniques for enhancing DGIS in a major way could be derived from the techniques found in the CONIT project at MIT. In this memo we outline a number of efforts that could be taken to expedite the ultimate incorporation of these new methodologies into DGIS (and, from there, to the world at large). In what follows we list a number of possibilities for enhancement and then outline some proposed efforts toward development of those possibilities.

B. PROPOSED ENHANCEMENTS

We may identify enhancements in 4 areas:

- (1) Problem and Search Preparation
- (2) Search Execution and Display
- (3) Search Evaluation and Modification
- (4) Support and General

As we discussed in our report, DGIS by itself offers very little in the way of assistance in any of these areas. On the other hand, the SearchMAESTRO (SM) module does provide some assistance; we shall make comparisons, then, primarily with this mode of DGIS utilization.

1. Problem and Search Preparation

1.1 Natural language keyword-stem search strategy

As analysis of the CONIT DGIS experiments demonstrated in support of much previous analysis, the CONIT methodology of creating a search strategy by performing a Boolean intersection of all-(subject)-fields truncation searches on the stems of keywords automatically extracted from users' natural language topic phrases is a very powerful tool for assisting searchers, especially inexperienced searchers in a multiple, heterogeneous database environment. In contrast, SM does very little to aid users develop effective search strategies, essentially providing only very basic examples easily understood by the casual enduser.

1.2 Index Term Browsing

CONIT permits a searcher to browse through the index terms from the current database. It also allows users to select terms from the index by tag numbers (instead of having to type out the full term) even in cases where the retrieval system itself does not provide those numbers. These features are not available in SM.

1.3 Common Command Language (CCL)

The CCL permits users to express their requests in one common form and have the system perform the appropriate translations to the currently connected retrieval system. Although there is now no CCL available in SM [*as of the date of this paper*], efforts in that direction are now being undertaken by both the DAITC Hypermedia Lab and Telebase for its EasyNet family (of which SM is one member)[*Note: Telebase incorporated CCL during the Summer of 1988*]. I feel that the CCL in CONIT is considerably more comprehensive than what has been achieved in either of these two efforts and our extensive experience and analysis of this issue, should prove beneficial to the further developments in either effort.

1.4 Search Definition and Delimiting

CONIT assists the user to prepare a formalized problem statement including a conceptualization based on a Boolean Topic Representation and a set of known and desired problem initial conditions and limits for such aspects as recall, costs (money and time), and document types. Besides helping the searcher be more clear about his needs, this formulation can be utilized later in retrieval evaluation and search modification, as we shall indicate below.

1.5 Database Selection

CONIT maintains a directory of all databases accessible on any of the retrieval systems and assists users in finding databases of relevance by leading them through a hierarchically arranged listing of the databases. Brief information about each database is immediately available in the listing and more detailed information may be requested from the retrieval systems themselves. Databases, either individually or in sets, may be selected by the user through indication of category or subcategory, full classification term, or character string or word at beginning of, or included in, CONIT full name or classification term or in alternate database names.

DTIC has plans for the development of a database directory and its utilization online in DGIS. SM has a form of database selection assistance in which the user selects answers to a series of menu options leading to the selection of a single database. Besides the above described full directory browsing scheme, CONIT has experimented with two techniques which provide a listing of databases ranked according to their likely relevance to the topic at hand. One such technique employs, as does SM, a series of menu formatted questions; ranking is accomplished through the application of a MYCIN-like formula operating on the user's answers. A second technique involves the searching of a multidisciplinary database (e.g., NTIS) by the keyword-stem techniques outlined in (1.1) and the cumulative transformation of classification terms found in the retrieved documents through a matrix whose rows are associated with classification terms, whose columns with databases, and elements with estimated relevance of each database to the topic expressed by the classification term.

1.6 Search History and Construction

Whereas SM primarily considers each search a separate monolithic entity to be discarded after its retrieval results have been initially reviewed, CONIT keeps track of all searches and search components and permits these to be operated on at any time during the session. Operations include:

(1) Display of any search results and documents (whether previously seen or not and whether still current or not in the retrieval system -- in the latter case regeneration of the "lost" search is automatically performed);

(2) Construction of new search strategies from old (note this can take place in CONIT itself while NOT connected to the retrieval system -- thus saving costs); and

(3) Search evaluations and running in (additional) multiple databases (see below for more details).

In addition to saving search history for a given session, CONIT has a search catalog facility which enables search strategies to be saved for a future session, even allowing sharing with other individuals or groups of searchers.

1.7 Comprehensive Search Operations

CONIT permits a full range of proximity specifications in searching on multiple words and terms (e.g., number of intervening words or inclusion within same [unspecified] field) whereas SM operates with the simple AND or strict adjacency specifications (i.e., just the ends of the proximity spectrum). As indicated below, there are plans to have CONIT be able to distinguish and specify and combination of subject-oriented fields for searching whereas SM is limited to a few variations of field searching.

2. Search Execution and Display

2.1 Component Search Recording

CONIT takes a compound (multi-word) search and breaks it up into its components which are then individually searched and may be separately reviewed or combined to form new compound searches (similar to the Dialog search steps facility).

2.2 Multiple database searching

CONIT permits a user to specify an ordered list of databases. Users may then request a search on one or all of the databases with a single request (even if the databases are on different systems). Note that the keyword-stem search methodology of (1.1) makes it feasible to get good results with the same search strategy across multiple, heterogeneous databases. The EasyNet "SCAN" feature is an initial start toward a true multiple database searching facility; SCAN apparently is limited to searching a few preselected sets of databases found in one system: Dialog. Note that in CONIT's "virtual system" approach users are not limited to framing their search for one (particular) system at a time.

2.3 Comprehensive Results Display

CONIT in its CCL, or through menus, allows searchers to select from four retrieval set output formats, two output presentations (online or offline), and any range of document records. By comparison, SM seems more limited in its flexibility. Also, we may note that CONIT's CCL may provide at least a methodological arguing point in considering the extension of the DAITC Hypermedia Lab.'s CCL:

(a) whether to adhere to the proposed NISO CCL standard in allowing the user to present arguments to the command in any order, and

(b) whether to try to make the command as much as possible common (or universal) across systems (as opposed to making it system dependent).

Also, the question of how to handle the interface between command and menu modes arises.

2.4 Connection Flexibility

CONIT permits user control of whether the remote retrieval system stays connected or not. One reason to want to stay connected is to avoid having to regenerate a search. (Of course, one must tradeoff quicker response time against a possible additional cost for longer retrieval system connect time.) CONIT does have a timer which warns of excessive connection time without interaction; such warnings could be replaced by automatic disconnection.

2.5 Sophisticated Connection Algorithm

CONT chooses among many alternate connection paths for a given database (e.g., X25 or dialout, Telenet or Tymnet, which system) not only on the basis of which system is supposed to have that database, but also in consideration of which system is currently connected (to avoid switching if not necessary) and which systems are scheduled to be available at the current time as well as making a dynamic choice depending on which paths have proven to be more reliable in the recent past (minutes and hours).

3. Search Evaluation and Modification

The techniques in this section include some that have been fully implemented and tested and some that have been designed but not yet implemented.

3.1 Recall Estimation

Three mechanisms for recall estimation have been implemented in CONT:

- o one based on an *a priori* comprehensiveness model (how many search terms and databases searched for each conceptual factor), and

- o two based on *a posteriori* relevance judgments as compared with, respectively, known and estimated numbers of relevant documents from *a priori* problem specifications (see 1.4).

A fourth, and potentially most definitive, estimation technique has been designed; it is based on sampled relevance judgments from purposely broadened searches.

3.2 Document Relevance Ranking

Designs have been made to enable the system to automatically rank documents on their expected relevance to the search topic. This technique is based on a model of how variations in search strategy relate to a metric of degree of association of a topic in relevance terms. Again, the problem conceptualization of (1.4) is central to the implementation of this technique.

3.3 Search Strategy Modification

In the first stage of the expert version of CONT which we have already implemented we have demonstrated how the user can be led through a series of menus to modify his search strategy to achieve designated goals (e.g., raise recall or precision). These menus include the selection from system-maintained lists of pertinent techniques for strategy modification. In the design for the full-fledged expert CONT we indicate how user relevance judgments on individual documents can lead to the automation of strategy modification technique selection, thus avoiding requiring the user to become involved with details of the techniques. Note that implementation of these enhancements has been assisted by the mechanism, already initiated, of providing modification <operators> that, when applied to existing searches, generate the searches that relate to the original searches in specified ways (e.g., change from a particular close proximity specification to an indicated looser one).

4. Support and General

We list below, without elaboration, a number of items related to user assistance and support or to general assistance system development issues.

4.1 General HELP, EXPLAIN, and ASSISTANCE facilities.

4.2 Integration of command and menu interface modalities.

4.3 Recording user comments and full or partial recording and printout of user's session. User memo creation and display.

4.4 Database and system cost rate display. Dynamic cost estimation in session. User and group cost accounting. Individual and group password control.

4.5 UNIX Shell command history and editing.

4.6 Production rule maintenance of protocol knowledge base.

4.7 Reconnect to dropped (or purposely suspended) process.

4.8 Workstation environment (currently being implemented).

4.9 Windowing, bit-mapped graphics, mouse/cursor input (planned).

C. ENHANCEMENT TASKS

In order to achieve the potential of a new, vastly higher level of effectiveness in search assistance inherent in the above enhancement techniques, DGIS needs to plan for the incorporation of such techniques into improved systems and the testing of these systems in order to determine what modifications are indicated for optimum performance. Based on its extensive, and often unique, experience in designing, developing, and testing these techniques, MIT is highly qualified for, and would be desirous of, assisting DGIS in carrying out such a program.

We would propose that a multistage effort be undertaken. In the first stage, lasting about 9 months, two kinds of activities could be undertaken:

- (1) Overall analysis and planning for the program, and
- (2) Initial implementation and testing tasks, particularly on the CONIT workstation system.

DGIS's ongoing planning activities are crucial in determining what new techniques on the technological forefront are worth pursuing and the means for coordinating and scheduling such efforts. DGIS has wisely chosen to move on a broad front of improvement activities including efforts involving the database directory, a common command language, other AI-based enhancement techniques, SearchMAESTRO, and CONIT testing, as well as a host of user training and other system testing and enhancements. MIT might act in a consultant capacity in support of DGIS's planning and coordination activities. In such a capacity we could explain in detail the techniques we have proposed -- including their rationales and detailed design and testing experience -- and we could suggest means by which such techniques could be incorporated into the various DGIS improvement programs with realistic scheduling and in a coordinated way.

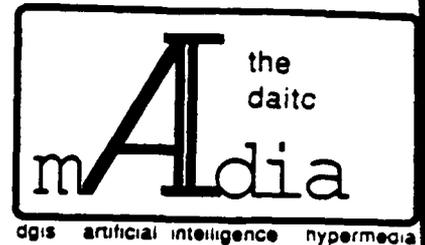
Initial implementation and testing of a workstation, UNIX/C version of CONIT can be achieved within stage one in two phases. Phase one would consist of converting CONIT's programs into a workable and demonstrable UNIX environment for portability to DGIS. In phase two the conversion would be completed, and reliable and smooth operation would be achieved so that operation and testing with users could be initiated. We have indicated in our report and in personal communications that we believe that this level of development would be sufficient to demonstrate major improvements in search assistance effectiveness over any existing retrieval assistance systems.

In the later half of stage one an assessment would be made of how to proceed to stage two. This could include the recommendation to proceed to a full-fledged version of expert CONIT, or to the alternate recommendation to incorporate some or all of the enhancement techniques in other DGIS improvement

programs, or some combination of these strategies. We believe we have shown in our report, and elsewhere, that we can make objective assessments of technologies -- including critical as well as praiseworthy judgments of our own system and techniques.

The level of effort required for the stage one tasks is in the order of four to six person months, primarily in the category of the principal investigator. The detailed scheduling in task area two, CONIT workstation implementation, depends on the degree to which we are successful in obtaining additional support, outside of DGIS, for our research program. In the worst case, a minimum level of CONIT operability (i.e., a phase one level) would be achievable; in more positive situations a more rapid and vigorous scheduling of expert CONIT implementation and testing would be achievable.

* Marcus, Richard S. **Experimental Evaluation of CONIT in DGIS Gateway Environment.** Lab. for Information and Decision Systems, Mass. Inst. of Tech., February 1988, LIDS-P-1746, AD-A204 721.



DGIS COMMON COMMAND LANGUAGE SYSTEM PROJECT

- & -

ARTIFICIAL INTELLIGENCE

DTIC's Common Command Language (CCL) project for the DGIS precipitated DTIC's entry into the Artificial Intelligence environment. DTIC began its initial prototypes in C language with DIALOG, BRS, NASA/RECON, and DROLS. These prototypes, developed in a third-generation syntactical language, brought to surface a number of issues relevant to information systems and the Common Command Language concept.

These issues concern both the user-machine interface and the CCL interface design. Creating a "standard" command language turns out to be a minor issue. The major issue concerns accomodating the operating characteristics and idiosyncracies of the individual systems.

The creation of a CCL is only one component of the "CCL-need" issue. A second component is creation of a CCL System that allows a user to search in unfamiliar systems without needing to know a system's operating characteristics. A third component is identifying the critical objectives that a CCL System is to serve. In the case of DGIS, the criteria for CCL objectives are the DGIS information processing operations.

The decision to convert to and continue our CCL development with Artificial Intelligence tools took place in April, 1987. PROLOG was chosen because it is a simple but powerful relational programming language based on the concept of programming in logic. Our effort merges PROLOG and C capabilities, to provide the DGIS user an AI-based searcher assistant interface. PROLOG programs will serve as the AI tool to handle the heterogeneous universe of diverse information systems. We are making use of knowledge base and blackboard architectures to achieve a universality across the information systems, and to create a human-machine interface that functions gracefully in the interests of the user.

We will apply a Common Command Language System (CCLS) to the major information systems: DROLS, NASA/RECON, DIALOG, BRS, and ORBIT. This will provide a standard manner of accessing 500+ databases. CCLS will then be extended to the multitude of databases both within and outside of the DoD community as DGIS acquires access to them.

As the laboratory resources expand with equipment and expertise, we will also expand into other AI-based programs in support of DGIS. These areas include:

Hypermedia Implementations
Information Processor Systems
Routine Generator Expert System
Human Foreign Language Interface
Portable Natural Language Interface
Diverse Database Query Expert System
Thesauri Integration for Expert Searching
Natural English Language Interface to UNIX
Numeric Information Query & Processing System
Information Analysis and ES Application System

The prime goal of the Hypermedia Lab. is to create AI-based programs that make the human-machine interface more human in the machine's responses.

**DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER (DAITC)
ALEXANDRIA, VIRGINIA**

OBJECT-ORIENTED PERSPECTIVE

--OO--

Object-Oriented systems are based on three concepts - *OBJECT*, *MESSAGE* (or method), and *CLASS* (or unit). Essentially, objects encapsulate a private data set which can be accessed or modified with the activation of their messages. Each object is an instance of a class. Classes are arranged hierarchially where designated properties of a class can be inherited by its subclasses.

A very simple example of an object-oriented system is as follows: Consider three objects, a, b, and c of type *Matrix-2*. The class *Matrix-2* is a subclass *Matrix* for matrices of size 2 by 2. To compute the expression: $a := b + c$, the user sends the message "add:c" to the object b, where "add" is a method inherited from a more general class *Matrix*. When b receives that message, the object creates a new object a that results from adding matrices.

Another example more closely related to us is found in the design of ProWINDOWS, a PROLOG window package for SUN workstations. The following is quoted from the ProWINDOWS manual:

>>> ProWINDOWS is an object-oriented programming system. In ProWINDOWS, an object is fully characterized by the task it can perform; nothing is revealed of its internal structure or of how the tasks are performed. To request that an object perform a task, one sends the object a message.

>>> The tasks that an object can perform are called behaviors. A window object, for instance, has an open behavior. If you send a window object an open message, the window invokes its open behavior and shows itself on the display.

>>> Each object is an instance of some class. A class is a way of grouping together a set of behaviors, so that the users may conveniently create new objects that have the same behaviors. Classes are objects in their own right, and as such have their own set of behaviors that you can invoke. Objects of class class include a list of all the behaviors their instances understand. So, you can ask a class what behaviors are defined for it, and thereby find out what messages a particular object can receive. <<<

The Object-Oriented (OO) approach originated with the Simula language and later was made popular by the Smalltalk language from Xerox PARC. Now most modern programming languages have some object oriented extensions: Object Pascal, Objective C, C++, Concurrent Prolog, Quintus ProTALK; LISP is flourished by a number of versions (Flavors, LOOPS, Common LOOPS, ..) and now the Common LISP Object System (CLOS) which is the standard being promoted by ANSI X3J13. Even ADA is claimed to have some object oriented supports by itself.

The benefits of OO far exceed any difficulties and draw a large following. The OO approach has these advantages:

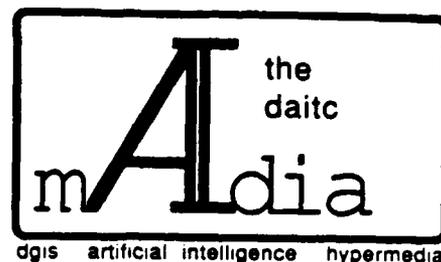
- It shortens the development time by facilitating concurrent work.
- It makes the product more flexible.
- It makes the product more comprehensible.
- It facilitates the correctness proofs for programs by providing a relatively simple statement of what we expect the module to do.

Beside programming languages, the influence of OO extends to the fields of knowledge representation, expert systems (most expert system shells now follow certain OO paradigms), distributed computing systems (i.e., distributed operating systems), CASE (Computer Aided Software Engineering), CAD/CAM, hypermedia, and database management systems.

There are now several commercial OODBMS on the market, such as Vbase, GBASE, and Staticc. These OODBMSs will allow the treatment of data in a non-linear fashion, as compared with traditional programming which handles data linearly and sequentially.

The prime goal of the Hypermedia Lab. is to create AI-based programs that make the human-machine interface more human in the machine's responses.

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER
ALEXANDRIA, VIRGINIA



GENETIC ALGORITHMS

"pssst - Darwin's back..."

Genetic Algorithms (GAs) are search algorithms based on the mechanics of natural selection and natural genetics. They combine survival of the fittest among string structures with a randomized information exchange to form a search algorithm. In every generation, a new set of artificial creatures (strings) is created using bits and pieces of the old; an occasional new part is tried for good measure. While randomized, GAs are no simple random walk. They efficiently exploit historical information to speculate on new search points with expected performance.

GAs have been developed by John Holland and his coworkers at University of Michigan. The main feature of GAs is robustness, i.e., the balance between efficiency and efficacy necessary for survival in many different environments. Furthermore, GAs inherit the capability for self-repair, self-guidance, and reproduction, in following the rules of biological systems; this barely exists in the most sophisticated artificial intelligence systems.

A Genetic Algorithm must have five components to solve a problem:

- [1] a chromosomal representation of solutions to the problem,
- [2] a way to create an initial population of solutions,
- [3] an evaluation function that plays the role of the environment, rating solutions in terms of their *fitness*,
- [4] genetic operators that alter the composition of children during reproduction, and
- [5] values for the parameters that the GA uses (population size, probability of applying genetic operators, etc.).

An example of a GA application is the problem of keyword assignment to facilitate document retrieval. A document in bibliographic databases receives multiple descriptions (each is a set of subject terms or keywords with associated weights). The problem is to provide a "best description" with respect to different inquirers seeking the same document in dissimilar ways. The genetic algorithm shown below is used to adjust these descriptions and provide a better means of getting documents to just those inquirers who find them useful according to their needs:

Cycle: Repeat -

- 1) For any particular document, measure the worth (i.e., performance or fitness) of each of its (fixed number of) descriptions. That is, determine how well each description serves in providing the document to just the right inquirer.
- 2) Replace the set of descriptions currently associated with that document:
 - a) Throw away its current set of descriptions.
 - b) Establish a new set of descriptions out of the set just discarded, using more parts of descriptions that had higher worth. Each of these newly created descriptions will likely be different from all descriptions in the just discarded set.

Result: Until - the set achieves your criteria.

The above algorithm is the basics of a GA. This is a process that attempts to mimic the patterns of genetics, promoting a population of descriptions built up of parts (genes) of its fittest member.

The prime goal of the Hypermedia Lab. is to create AI-based programs that make the human-machine interface more human in the machine's responses.

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER (DAITC)
ALEXANDRIA, VIRGINIA

CONNECTIONIST MODELS (or NEURAL NETWORKS)

"--flash s(y)nap(se)--"

Artificial Intelligence (AI) has traditionally followed the top-down approach of solving a task by breaking it into smaller processes until one can actually build them, either in hardware or software. This approach was logic-based and it emphasized a well-defined *representation* and an *inference system* which was generally driven by some *heuristics*. A lot of successes had been reported using the above approach until the early 1980s when it became clear that new paradigms were needed to overcome some serious limitations:

- * The problem of knowledge acquisition becomes a strong reality for large expert systems. There was also the interest in providing AI systems with a learning capability to adapt to the environment when it is changed, with the goal of decreasing the brittleness of the system.

- * The famous frame problem (roughly, deciding what to change and what not to change when a robot's model of the world changes) was considerable, but that changed in light of new developments in *non-monotonic logic*.

- * Computer vision was functionally constrained under the top-down approach. There was constant difficulty in making all the connections needed for computer vision to achieve life-like functionality.

Connectionist Models (CMs) or Neural Networks (NNs) are enthusiastically advocated as the alternative approach to address the above limitations. For example, under the neural network concept computer vision has flourished and is exploiting new special-purpose hardwares that are neurally inspired. Neuroscience also provides steady advances in knowledge and technology in nervous system studies.

Basically CMs go bottom up. You begin with the parts that you understand (simple neurons). Then you work upwards in complexity, to find out how to interconnect those units to produce the larger scale system that can solve the real-world problem. The bottom-up paradigm circulated in academic circles in the 1960s, but it did not start to be viable until two critical elements appeared: the state of the art of hardware toward parallelism, and our understanding of how to embody ideas in machines (i.e., knowledge representation).

A number of variations of neural networks have been introduced. Their differences are in the number of layers of neurons, the connections between neurons, and the output functions of the neurons. The major points of such neural networks are:

- * CMs integrate *all* available evidence. Some of these pieces are irrelevant or only weakly relevant for the output. Contrasted with the logic-based approach, a CM does not have to find necessary and sufficient conditions; statistical correlations between actions and the facts of the world can be represented as weighted connections. Connectionist reasoning is *evidential* rather than *logical*.

- * In a CM a large number of neurons (or nodes, or modules) contribute toward a connectionist decision. Thus, it is not necessary for all the neurons to be active in order to make a correct decision; some variations of values can be tolerated. This fault-tolerance characteristic is a strong point of CM. The CM does not break when it is faced with an unknown environment, but rather degrades gradually and gracefully.

- * In order for AI to deliver solutions to real world problems, it must be demonstrated with large problems. The dramatic strides in massive parallel computer architecture (e.g., Connection Machine, Transputers) provide the hardwares that are well-suited to CM, and *vice versa* much CM research has inspired several neuro-computers.

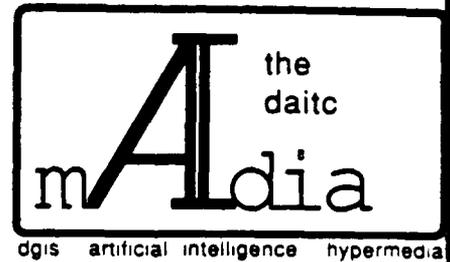
- * CMs allow much easier integration of modules than the traditional AI programming environment with symbolic processing and heuristics. Modules in CMs are interconnected by simple weighted links. Messages themselves can also be very simple (e.g., numerical activation levels, markers). In some cases, link weights can be automatically generated based on statistical or structural analysis.

- * Learning is the most exciting aspect of CMs. A number of methods such as back propagation error learning and the Boltzmann machine are quite effective. CMs have exhibited non-trivial learning; they are able to self-organize, given only examples as inputs.

The prime goal of the Hypermedia Lab. is to create AI-based programs that make the human-machine interface more human in the machine's responses.

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER (DAITC)

ALEXANDRIA, VIRGINIA



C LANGUAGE INTEGRATED PRODUCTION SYSTEM (CLIPS)

An ES shell from NASA

The DAITC Hypermedia Lab acquired in Nov 88 an expert system shell called CLIPS (C Language Integrated Production System). This ES shell was developed at the Artificial Intelligence Section, Lyndon B. Johnson Space Center, National Aeronautics and Space Administration (NASA), Houston, TX. In NASA, CLIPS shell programs are running in VAX/VMS, MacIntosh, and IBM PC environments.

As indicated by the name, CLIPS is written in C, but makes use of that high-level third generation language to create an ES tool environment. CLIPS therefore contains the basic ES elements of a fact-list, or a global memory for data, a knowledge-base, containing all the rules, and an inference engine, which controls overall execution. CLIPS, therefore, is a data-driven program where the facts are the data that stimulate execution, as compared with procedural languages such as PASCAL, ADA, BASIC, FORTRAN, and straight C. One consequence of this environment is that in CLIPS rules are activated in parallel, while the statements of the other languages are sequential, line by line.

The CLIPS major characteristics are:

- built in C
- can turn rules into C (if desired)
- a subset of ART (an ES Shell by Inference Corp.)
- nice integrated environment for Mac
- extremely portable
- fast
- low memory requirements (relative to other systems)
- full source code is available

The Hypermedia and DGIS Labs have gotten CLIPS running in the UNIX/C environment with local enhancements, totally integrated with the DAITC UNIX environment. Its availability on the Sun, VAX/BSD, and Pyramid is now possible.

The Hypermedia Lab foresees embedding CLIPS-based programs in several on-going developments, including, for example, the Common Command Language System (CCLS) for DGIS, and the Hypermedia Lab-produced Hypertext program on the VAX. A highly foreseeable near-term application is a NAM Generator ES (See AD-A181 103, *Toward an Artificial Intelligence Environment for DTIC: Proposed Tasks*, ..., May 87, p.4, para 5, sub-para.: "Connection Agent Generator"). Demo CLIPS ESs include a Wine Selection ES and an Automobile Diagnostic ES, available on the Pyramid and the VAX.

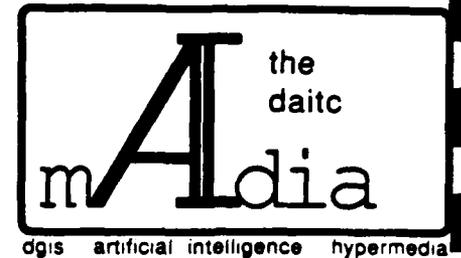
Because of the features listed above, CLIPS is a very powerful tool that can provide AI-embedded program capability to C software, or can be used as a CLIPS shell-based stand-alone application. Thus, CLIPS-based programs developed at the DAITC could provide AI-based programs that are highly exportable directly to other UNIX environments.

Also, although the Lab. has acquired AI software from the public domain, primarily through downloading, CLIPS is the first AI software obtained through interagency request.

The descriptions of CLIPS above are taken from the CLIPS documentation provided by NASA.

The prime goal of the Hypermedia Lab. is to create AI-based programs that make the human-machine interface more human in the machine's responses.

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER
ALEXANDRIA, VIRGINIA



*** PROLOG ***

An Artificial Intelligence Programming Tool

The founding of the Hypermedia Lab.

Per the *Catalogue of Artificial Intelligence*, Alan Bundy, Springer-Verlag 1986:

PROLOG is A simple but powerful and practical programming language based on the idea of programming in logic. PROLOG programs may be viewed as logical clauses and the interpreter as an efficient resolution theorem prover. PROLOG may be looked on as an extension of LISP in that it provides as primitives pattern directed procedure invocation and non-determinism (backtracking). It provides general recursive (tree-like) data structures which are accessed by pattern-matching rather than by explicit selector functions. There are no destructive operations on these data structures, but structures may contain empty slots (uninstantiated logical variables) which can be filled in later. There is also an assertional database which is used for relatively long-lived or permanent data.

The Hypermedia Lab's entry into Artificial Intelligence (AI) came from its experiences in the development of the DGIS Common Command Language in UNIX/C on the DGIS VAX 11/780 during 1986-87. After assessing the rigidity of 3rd generation programming (C, with UNIX utilities), coupled with the requirement for universality and a more human-like human-machine interface, we decided to make the jump to AI in April, 1987. A review of the AI tools relative to the needs of CCL led to the decision to restructure the DGIS CCL in PROLOG, still in a mainframe environment (VAX 11/780). The reasons were as follows:

The Reversibility of PROLOG: Logic programming concerns the relationships of objects (or terms). In determining the truth of relations, we can have reversibility in programs, *that is*, one can write a program and have its inverse for free (with some restrictions). This feature provides a tremendous advantage to CCL, for example, in that CCL concerns the translation of a multiplicity of diverse search languages.

The Database Capability of PROLOG: PROLOG has its own internal database capability. This feature allows programs to manipulate codes as relations that can be asserted or deleted. The feature can also be extended to external databases (*e.g.*, via a RDBMS), to achieve the flexibility of storing knowledge in both PROLOG internal databases and traditional external databases. We thus take advantage of database technology for performance and ease of use. (Commercial databases are equipped with a number of utilities to maintain and access their data.)

The Separation of Logic and Control: This separation is encouraged in PROLOG. PROLOG programs can be considered an amalgamation of rules and facts. They are governed by the default execution control of the PROLOG language, *i.e.*, left to right and top to bottom. This control can be easily supplemented or replaced by more powerful meta-rules, coded also in PROLOG.

Object Inheritance & Message Passing: These are two powerful programming features in object language methodology. Object inheritance and message passing are very easily implemented and embedded within PROLOG. Object language methodology has been used in CCL, in conjunction with the above capabilities of logic programming language.

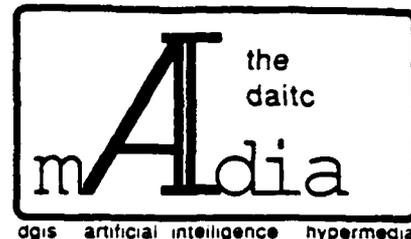
Per the *Catalogue/Bundy*:

Since it's adoption as a base language for the Japanese Fifth Generation Project, PROLOG has exhibited a remarkable rise in popularity. Commercial and/or research implementations are now available for almost every conceivable architecture, ranging from home micros to large scale mainframes.

For a brief on CCL and AI, see Hypermedia Lab. Program Note 3, May 88.

The prime goal of the Hypermedia Lab. is to create AI-based programs that make the human-machine interface more human in the machine's responses.

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER, ALEXANDRIA, VIRGINIA



QUINTUS PROLOG * A Vendor Package Version
"The Quintessential PROLOG"

Quintus Prolog is a PROLOG commercial implementation by Quintus Computer Systems, Inc., Mountain View, California. Quintus was started by a group of researchers that wrote the world's first PROLOG compiler at University of Edinburgh, Scotland.

The Quintus implementation of PROLOG is fast and efficient and it is available on a number of UNIX computers and the DEC VMS operating system. We selected Quintus in 1987 after a period of evaluation and testing. A version of Quintus Prolog was installed on the VAX 11/780, and it has been used to implement the DGIS Common Command Language System (CCLS) and hypertext systems. In October 1988 we obtained on loan Quintus ProWINDOWS for evaluation. ProWindows is an object-oriented PROLOG-based package for Sun workstations. It could be used for implementing a bitmapped user-interface to CCLS, hypermedia, and Neural Network applications. We observed the ease of programming and expressibility provided by this product that is not generally provided elsewhere.

The major features of Quintus Prolog include:

- * A fast, efficient implementation.
- * Easy program entry and debugging using the Emacs text editor interface.
- * Interface with programs and modules written in other languages, such as C.
- * High degree of compatibility with the industry standard of PROLOG (Dec-10 Prolog).
- * On the Sun workstation Quintus provides ProWINDOWS, an object-oriented programming package which enables programmers using Quintus Prolog to quickly and easily create window-based user interfaces for their PROLOG application programs.

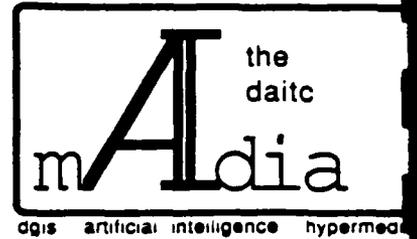
With Quintus Prolog on the VAX, the Hypermedia Lab, in addition to CCLS, has created several PROLOG-based programs. One is an implementation of hypertext processes, available as a shared computing resource (*vis-a-vis* a PC standalone package). Those programs in essence have changed the VAX to a hyperVAX system, and put the PROLOG-based programs into a distributed programs environment. In this application we structure the PROLOG programs as processes separate from the VAX's C frontend process. The communication protocol between PROLOG and C to send requests and retrieve answers is provided by Quintus Prolog. This facility is based on standard BSD UNIX sockets and it is useful for several reasons:

- * It allows PROLOG to function as a back-end component of another system written in another language.
- * It takes advantage of the computing power of a network of computers by allowing the construction of a set of cooperating PROLOG processes running on multiple machines.
- * It permits PROLOG to control distributed processes.

A strong advantage to acquiring a commercial package is the inclusion of vendor maintenance and support, such as software updates, documentation updates, and vendor conferral support. This maintenance is especially critical for leading-edge technologies, such as PROLOG as an Artificial Intelligence tool, as they continue to evolve.

The prime goal of the Hypermedia Lab. is to create AI-based programs that make the human-machine interface more human in the machine's responses.

**DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER
ALEXANDRIA, VIRGINIA**



NU-PROLOG * An Example of a University Version

"What's NU?"

NU-PROLOG is an example of a university PROLOG development application, *vis-a-vis* a commercial development. Whereas a university version is a research development made available mostly for non-commercial uses, a commercial version is vended as a frozen version, to include options for maintenance, training, and later updates. A second major difference between university and vendor versions is that a university version usually is relatively low cost, does not include maintenance and training, but reflects the state-of-the-art of the technology.

NU-PROLOG is a development of the Machine Intelligence Project of the University of Melbourne, Australia. The Hypermedia Lab's interest in this particular version stems from its capability to be installed on the Pyramid 98x, the DTIC/DAITC operational DGIS machine; contains features reflecting the latest advances in the logic programming field; and is compatible with Quintus PROLOG, the vendor package on the DAITC VAX 11/780. At the time of the transition from the VAX to the Pyramid in August, 1988, Quintus had not modified a package for the Pyramid. A review of PROLOG packages during May-June preceeding brought up NU-PROLOG.

NU-PROLOG succeeds an earlier MU-PROLOG, and moves PROLOG closer to the ideals of Logic Programming by allowing the user to program in a style closer to first order logic. It provides substantial performance gains over the earlier interpreted systems such as MU-PROLOG, with the following features:

- * Compiles PROLOG programs into machine code for an enhanced version of the Warren abstract machine.
- * Incorporates a database system based on superimposed codeword indexing which can store general PROLOG terms in external databases for fast retrieval; the database system makes use of the superjoin algorithm to perform efficient join operations.
- * Uses "when" declarations (successor to "wait") to control execution of programs according to the availability of data.
- * Implements a large set of built-in predicates, including many Quintus PROLOG predicates; most DEC-10/Edinburgh/MU-PROLOG library predicates are available through compatibility libraries.

The NU-PROLOG system contains the following major components:

- * "nc", the NU-PROLOG compiler.
- * "np", a simple interpreter-style interface which implements the standard Edinburgh PROLOG style debugging facilities and has a sophisticated query language for accessing external database predicates.
- * "nac", a program for adding control information to programs written in a purely logical style.
- * "nit", a program for reporting common errors in programs (*cf.* UNIX/C's "lint").

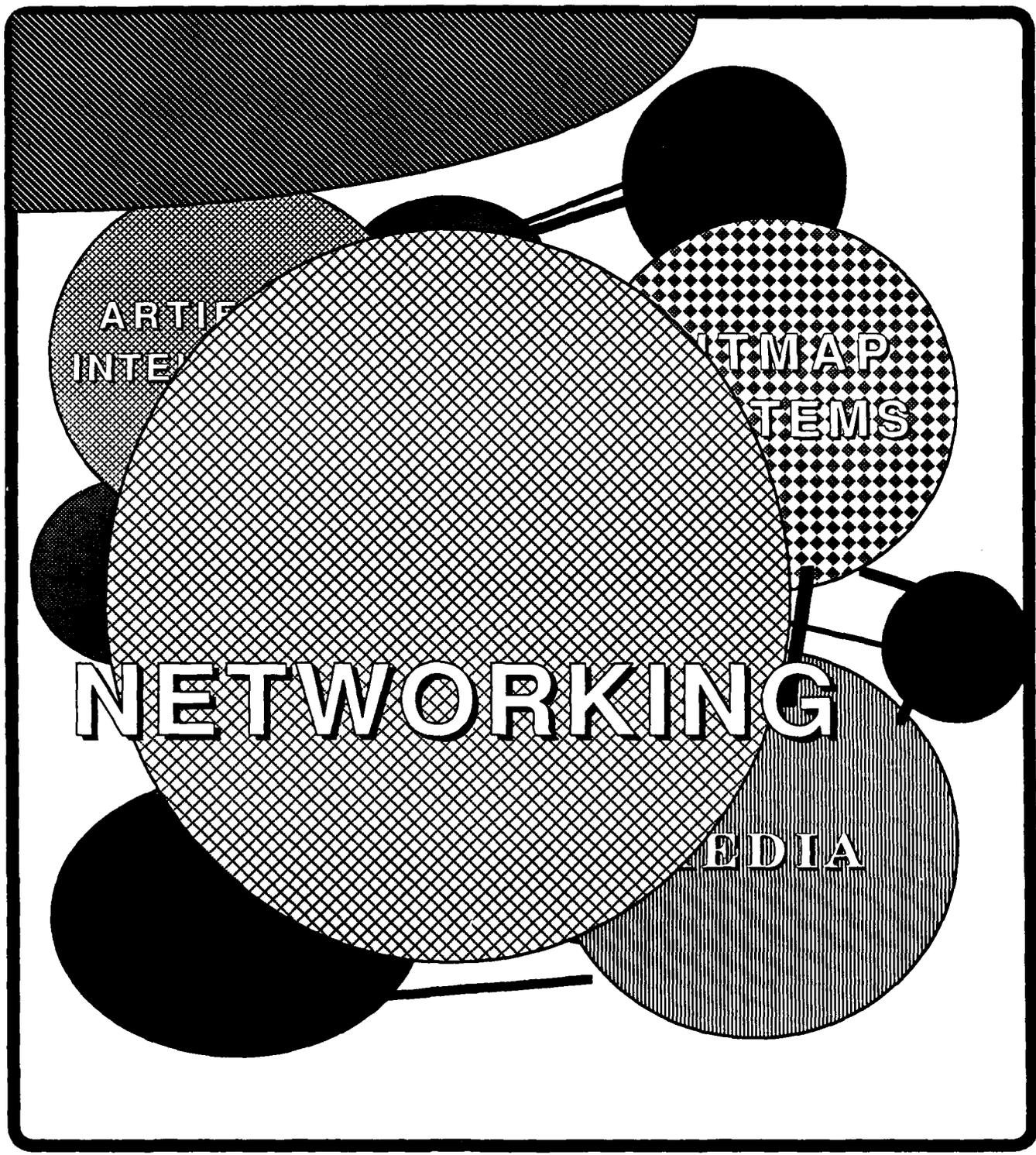
The NU-PROLOG 1.3 release includes an improved interpreter and debugger which understand "when" declarations; floating point arithmetic; an interface to foreign functions on many machines, and the usual large collection of bug fixes. NU-PROLOG runs under UNIX System V and Berkeley BSD UNIX 4.#. It has been implemented on VAX 11/780, Pyramid 98x, Elxsi 6400, Perkin Elmer 3240, Sun workstations, and Integrated Solutions Workstations. The Hypermedia Lab has or communicates with the VAX, Pyramid, and Suns.

NU-PROLOG descriptive information above has been picked up verbatim from information provided by the University of Melbourne.

The prime goal of the Hypermedia Lab. is to create AI-based programs that make the human-machine interface more human in the machine's responses.

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER

ALEXANDRIA, VIRGINIA



DAITC HYPERMEDIA LABORATORY

GLOBAL SCIENTIFIC AND TECHNICAL INFORMATION NETWORK

Gladys A. Cotter
Director
Defense Applied Information Technology Center

The section following is reproduced from:

Cotter, G. A. "Global Scientific and Technical Information Network;" *Online Information 88, 12th International Online Information Meeting, London, 6-8 December 1988*, pp. 611-618. Available as AD-A201 902.

Background reference for this paper is:

The Scientific and Technical Information Network (STINET): Foundation for Evolution.
Gladys A. Cotter, Defense Technical Information Center, September 1987, AD-A189 750.
Also published as "Information Retrieval Systems Evolve -- Advances for Easier and More Successful Use;" *AGARD Conference: Barriers to Information Transfer and Approaches Toward Their Reduction*, NATO/Advisory Group for Aerospace Research & Development (AGARD), AGARD-CPP-430, Paper 5, September 1987.

GLOBAL SCIENTIFIC AND TECHNICAL INFORMATION NETWORK

G.A. Cotter, Defense Applied Information Technology Center, USA

Keywords: *Networks, Information Technology, Gateways, International, Interfaces*

Abstract: This paper describes the foundations for a global scientific and technical information network which is currently under development at the Defense Applied Information Technology Center (DAITC). The paper highlights advances in information technology which are creating an environment where networking and interoperability strategies are possible on a global basis. Applied information technology modules which have been created to support the Global Scientific and Technical Information Network are described. The network is purposefully designed to incorporate both end-users and their information management intermediaries in a complementary manner making them resolute partners in the work and its rewards.

1. NETWORK OVERVIEW

Scientific and Technical Information (STI) is being produced, enhanced and stored around the globe. Single countries in some cases are acknowledged leaders in select scientific and technical disciplines. But in today's interdisciplinary and interrelated environment, it is almost impossible to cordon off STI needs into concise, neatly bounded subject specialties. Most scientist and engineers need access to STI on a global basis in order to maintain a state-of-the-art awareness. Global economies dictate that every effort be made to reduce unnecessary product and service development costs. This makes it necessary to ferret out and exploit to its full potential globally scattered STI which can reduce these costs.

The United States (U.S.) Defense Technical Information Center (DTIC) is laying the foundations for a global scientific and technical information network (STINET) to service its user community. The purpose of the network is to facilitate Department of Defense (DoD) access to scientific and technical information relevant to DoD mission areas. Key elements of the STINET are DoD managers and scientists; DoD Libraries and Information Analysis Centers (IACs); DoD Laboratories; DTIC; and other DoD, federal, commercial and international databases, systems and sources of STI. DTIC is tasked to provide a focus for molding the STINET, to establish interoperability among the various network components, and to coordinate network evolution. DTIC is to achieve this vision by promoting resource sharing and cooperative efforts and through investigation, experimentation, and application of advanced information science and technology.

Development of a STINET is a huge and challenging task which must be accomplished within existing financial and personnel constraints. In order to make visible progress towards the STINET, DTIC had to carefully identify network requirements and evaluate the currently available technology in which applied research could be invested in both the short term and the long term to meet these requirements. Next, DTIC had to settle on an agenda of purposeful steps that would close in on the ultimate goal. Successful development and implementation of the network depends on devising a "doable" plan of work with room for deviation when opportunities for technological acceleration become apparent.

The following elements are required for a successful network:

- o interoperability and interconnection among geographically disperse systems
- o tools such as pointers and menus, to help locate information on a global basis
- o standardized command functions
- o compatible, multifunctional, flexible software for installation at network nodes
- o interconnection with diverse sources of information including government, commercial and international
- o interconnection with diverse types of information including numeric and factual
- o selectivity and data analysis routines
- o improved delivery systems
- o integration of databases and people bases

In addition, DTIC must ensure that the network design incorporates both information end-users (scientists, engineers and their information specialists, etc.) in a complimentary manner. In order for the network to evolve, these two groups need to become resolute partners in the exploration of how new technology can be applied in an actual information service environment.

2. NETWORK DEVELOPMENT

Prototype Approach

The technology needed to meet the requirements of the STINET was at various stages of readiness. Rather than attempting to build an operational system immediately, DTIC decided to develop prototype systems. The prototype approach provided a means for rapidly integrating new technology components into the network for test and evaluation by actual users with functional requirements. It allows a low level of financial-investment and an opportunity to work with test user groups to add, revise or delete capabilities before introducing them into a production environment with operational users. DTIC selected several key and promising areas in which to invest prototype development efforts. The goal of these efforts was to provide the basic foundation for, and functional capabilities of, the STINET.

Having settled on the prototype development model, DTIC moved to establish a suitable environment for this activity. In partnership with several other DoD organizations who had similar needs, the Defense Applied Information Technology Center (DAITC) was established. The DAITC's mission is to facilitate the prototype development, application and introduction of new information technology. The DAITC is organized into technology laboratories which concentrate in the following areas:

- Interoperability/Networking
- High Density Information Systems
- Artificial Intelligence
- Video/Voice Systems
- Hypermedia
- Database/Text Search

Each laboratory is equipped with hardware and software which is available for use in prototype development, test and evaluation. The DAITC also has a powerful computer and communications network which allows interconnection with international communication networks and resources.

DTIC's prototype efforts are primarily being accomplished at the DAITC. The DoD Gateway Information System, the Integrated Bibliographic Information System and the SearchMAESTRO system (described below) provide examples of prototype systems which are being developed at the DAITC as the foundation for STINET.

The DoD Gateway Information System (DGIS)

The DoD Gateway Information system (DGIS) is an intelligent gateway system which provides information access and analysis. The capabilities goal was to link people, information

services, and computers pertinent to the STINET. The technology embodied in the DGIS will provide the key menus, pointers, interoperability, and interconnection within the STINET. To accomplish this, DGIS will have to function as an electronic switch, a communications interface, and a transaction controller. DGIS will require a variety of alternatives, tailored to different user types and needs, for obtaining and distributing information.

DGIS focuses on streamlining the information retrieval and analysis process. This is accomplished by placing the user at the center of a vast information universe consisting of people bases and databases and providing the user with the navigational tools required to pinpoint and arrive at his destination. In terms of databases, the DGIS is designed to provide users with answers to the questions.

WHAT RELEVANT DATABASES EXIST?

HOW DO I ACCESS THEM?

HOW DO I RETRIEVE INFORMATION FROM THEM?

HOW DO I MANIPULATE THE RETRIEVED INFORMATION?

The DGIS provides a single, easy-to-use interface for identifying, accessing, interrogating, and post-processing information from numerous databases relevant to DoD information needs.

In terms of people bases, the DGIS is designed to answer the questions:

WHAT EXPERTISE IS AVAILABLE ON THE NETWORK?

HOW DO I COMMUNICATE WITH EXPERTS?

HOW DO I SHARE INFORMATION WITH COLLEAGUES?

The DGIS acts as an integrated information system which allows human experts, information users, and information resources to exist and interact in harmony.

Development of DGIS is a multi-year, multi-task project. A prototype system has been developed which was designed for a user community including both intermediaries and end-users. Databases accessed are federal, commercial and international. In addition to large, well-known databases and systems, many small, specialized DoD databases will eventually be part of the DGIS.

The basic components of the system are:

A DIRECTORY OF RESOURCES, SUBJECT SEARCHABLE

A COMMON METHOD FOR ACCESSING AND SEARCHING DIVERSE DATABASES

TOOLS FOR DOWNLOADING AND POST-PROCESSING DATA

TOOLS FOR COMMUNICATING WITH A NETWORK OF EXPERTS AND COLLEAGUES

These are described below.

Directory of Resources

The Directory of Resources will include subdirectories with references to databases, people, and computing resources. In the first phase, the Directory of Databases is being developed using the INGRES relational database management system. The Directory of Databases contains information on the content, scope, and availability of selected databases. The Directory is subject-searchable; upon entering a topic of interest, the user is provided with a list of appropriate databases.

The Directory of People will contain references to experts in subject areas and information retrieval techniques who may be contacted via the network. The Directory of Computing Resources will contain references to computing resources, such as supercomputers which can be used for data analysis and modeling, available through the network.

Interfaces for Searching Diverse Databases

One of the primary goals of DGIS is to relieve the user of the need to learn and master separate commands and protocols for each database. As mentioned earlier, the DGIS target user community includes both end-users and intermediaries. DTIC found that end-user and intermediary interface needs are very different. (An expansive natural language interface requiring artificial intelligence applications appealed to both populations, but could not be accomplished with existing technology in the short term.)

A dual approach was adopted for the interface design, incorporating separate strategies for intermediaries and end-users. For the intermediary, a common command language is being developed. To satisfy the end-user, the EasyNet database searching service has been integrated into DGIS under the name SearchMAESTRO. SearchMAESTRO is a menu-driven database front-end which provides access to over 900 commercial databases. SearchMAESTRO access is now an option within DGIS and is described in greater detail in a later section of this paper.

Post-Processing

Information retrieved from databases often requires analysis or post-processing in order to become useful to the researcher. A library of post-processing routines for numeric and bibliographic are available in DGIS. In order to post-process data, the user downloads it into a file on DGIS, translates the data into a common format, and calls up one of several available post-processing routines.

People Bases and Databases

The goal in obtaining information is to acquire knowledge. Much of the information we need resides in the minds of human experts. Therefore, the DGIS has been designed to allow interaction among people, hence, the concept of people bases as well as databases.

Through the DGIS, users will be able to identify and communicate with experts and colleagues and to connect to information resources. As a first step, we have focused on providing the technology to allow such interaction.

Accessing a database/system is accomplished using the CONNECT command. This command provides users with automatic access to information resources. Users do not have to know telephone numbers, Defense Data Network (DDN) locations, passwords, access protocols or logout protocols. The user enters the CONNECT command and a data resources name. DGIS then establishes a connection to the resource and logs the user in. DGIS uses TYMNET, TELENET, DDN, and commercial telephone lines to establish connections.

The CONNECT command can be used to access information centers worldwide. To be eligible to use the CONNECT command to access a resource, a DGIS user establishes a deposit account with DTIC so they can pay for usage, and the DGIS System Administrator grants them connect permission.

Several mechanisms are available for interconnecting people. An electronic mail service is available twenty-four hours a day. Standard electronic mail features such as send, receive, answer and forward are incorporated. Mail messages can be sent simultaneously to multiple addresses and to every member of pre-established mail groups with lengthy documents attached if needed.

In addition, electronic mail can be used to send information downloaded from a database and placed in a file. A user who does not want to do his own database searching can send a search request to an information specialist without leaving his desk. The information specialist can perform the search, download the results to a file, and send the file to the user. Since the data is stored in a file, both the information specialist and the end-user can use post-processing routines for manipulating and analyzing the data.

The LINK command allows two or more users at different locations to link their terminals so that they view the same data display. All users have control over the display and can issue commands. Through the LINK command information specialists and end-users can together perform interactive database searches. The end-user benefits from the specialist's expertise while the specialist benefits from the end-user's immediate feedback.

Prototype and Beyond

A prototype DGIS has been developed and is currently undergoing test and evaluation within the DoD community. There are currently 150 users testing the system. The DGIS prototype is running on a VAX 11/780 using the UNIX operating system, the INGRES database management system, and a PROLOG interpreter package. The DGIS software is being ported to a Pyramid 98X, an ELXSI 6800, a Gould 6050, and Sun Workstations for

benchmarking and performance evaluation. Based on the results of the performance evaluation, a hardware configuration for a production system will be acquired. The production configuration may consist of several machines networked together. For example, the common command language and post-processing routines could be isolated on a back-end machine. DTIC plans to stabilize a version of the DGIS and offer it as a standard DTIC service in October 1988. Prototype development will continue on a separate development machine, and enhanced versions of the DGIS will be made operational at selected intervals.

The DGIS was developed in prototype as an unclassified, minicomputer-based centralized gateway system. As we move this version into operation, we plan to begin development of a distributed, clustered gateway network. Gateway nodes in the network would be made up of centralized, multi-user configurations and intelligent workstations distributed to users. A selection of gateway capabilities would be available on personal workstations. It makes sense, for example, to have automated connection routines, common command languages, and post-processing routines for frequently-used databases available on a personal workstation. The centralized, mini-based node would be utilized to identify and search infrequently-used resources and for post-processing volumes of data which are beyond the capacity of the personal workstation. Clustered gateway nodes consisting of a centralized gateway processor and personal workstations could be based on geographic, organizational, or subject boundaries. The clustered gateway nodes would interconnect and route users to appropriate nodes when necessary. Development of a classified gateway system is also underway.

Integrated Bibliographic Information System (IBIS)

Development and implementation of the DGIS allows DTIC users to connect to, search and analyze data retrieved from diverse unclassified database services in the federal, commercial and international sectors. Development, as planned, of a classified version of the DGIS will make its reach almost limitless. But what the DGIS has not provided is tools for the development of local databases or catalogs of holdings and tools for local collection management.

The DGIS is designed as an intelligent switching mechanism. The resources targeted by the DGIS are already online. The Directory of Resources is the only database central to and created and maintained on the DGIS. This is a basic design philosophy, not to be altered for fear of deflecting the DGIS from its primary focus -- that of being a gateway.

But there was a need to provide a vehicle to automate and manage local information collections which are manually maintained, very valuable, and very difficult for non-local personnel to use. This need is acutely felt by the DoD library community, a key component of STINET.

Therefore, DTIC initiated development of a library automation system responsive to the networking and local collection management needs of DoD libraries. The system would support centralized resource sharing while allowing local processing flexibility. The objective was to permit DoD libraries to make maximum use of existing information, organize this information to meet local needs, and selectively share newly-generated information with other members of the community. The system designed to accomplish this would have to integrate local control for local collection management functions (reference, cataloging, and circulation) with access to the external resources required for reference, shared cataloging, and other network requirements.

The Defense Nuclear Agency (DNA) was selected as the prototype site for the IBIS. The DNA library had a collection of more than 100,000 holdings and required a system supporting multiple users. The IBIS prototype was implemented at DNA on a VAX 11/750 minicomputer.

Scaling Down. The hardware configuration required to support the prototype IBIS at DNA, a minicomputer-based system, was not economical for smaller DoD technical libraries with collection sizes ranging from 5,000 to 75,000 items. A microcomputer-based system was more appropriate for the lower transaction volumes and smaller operating budgets associated with these libraries. Therefore, DTIC initiated an effort to identify and isolate the special

requirements of these smaller DoD libraries and accommodate them. A software survey was performed to identify packages which were suitable for servicing these smaller libraries and, at the same time, were compatible with the gateway software. The U.S. Army Training and Doctrine Command (TRADOC) library was selected as the prototype site for the microcomputer-based version of the IBIS.

Future Plans. The results and experience gained during the test and evaluation of the DNA and TRADOC prototypes have been used to develop the specifications for a competitive procurement of a production system from a commercial source. The production system will be available for purchase and installation by any library on the STINET in 1989. As a result of the dual approach -- small and large libraries -- an IBIS product line which can meet the needs of any DoD library, regardless of collection size and transaction volume will be available.

The IBIS is the first of DTIC's Local Automation Models and will make network-compatible software available for local installation and use. The IBIS is tailored for bibliographic information, and its target community is DoD libraries. Later, Local Automation Model product lines can be tailored for other user communities and different types of data. For example, models tailored for numeric data may be a requirement for some communities. Depending on the software selected for the production IBIS and the non-bibliographic user requirements identified, it may be possible to modify the IBIS software to meet additional needs. If not, DTIC can develop additional network-compatible local automation models specifically for an identified defense requirement.

Successful deployment of the DGIS and IBIS will provide DTIC with a powerful product/service line for the STINET. Through STINET, users would have a mechanism for interconnecting with globally available information systems, computing resources and people. They would also have a network-compatible vehicle for automating and managing local information, selectively sharing that information with other members of the network, and analyzing information from local and remote resources.

SearchMAESTRO

Both DGIS and IBIS were designed to deliver the power and utility required by a broad section of the targeted STINET users. But part of this user community did not require the full power of the DGIS and the IBIS. This group consisted of end-users who wanted to do some of their own information gathering. The users in this market segment had the requirement to scan literature and locate relevant items in their area of interest. They needed an interface that would provide easy access to a variety of databases. In some cases, the users would not be familiar with existing databases, so they required a system that would select a database for them and guide them through the search process. The information needs of these users tended towards fact retrieval -- such as the latest production statistics for a manufacturing company -- or scanning for relevant items -- such as what newspaper articles have been written on a particular subject. Their need was for some information relevant to the subject rather than an exhaustive search of the subject. We refer to them as casual end-users.

For this segment of the community, DTIC introduced SearchMAESTRO. As mentioned earlier, SearchMAESTRO provides access to over 900 databases. SearchMAESTRO can be used directly or through the DGIS. Users who access SearchMAESTRO directly can access databases via a simple-to-use interface. This interface eliminates the need to learn unique database command languages and search techniques. The user has two modes of operation from which to choose. In the first mode, SearchMAESTRO leads the user through a series of questions and answers and select the database for him.

In the second mode, the user can select the database he wants and use the SearchMAESTRO interface for executing the search. With either mode, the user can view search results on the screen, print, or save the results using local equipment. The main reason for accessing SearchMAESTRO through the DGIS is to take advantage of the post-processing and electronic mail utilities. The users must always observe fair use practices when dealing with copyrighted material.

The unclassified portions of the Technical Reports and the Work Unit Information system files of DTIC's Defense Research Development, Test and Evaluation Online system (DROLS) are being made available to registered users through SearchMAESTRO.

Any time man and computer meet, a diversity of problems arise which can best be handled through human intervention. Therefore, SearchMAESTRO provides online user assistance through a function called "SOS" for "Save Our Search". At any point during a SearchMAESTRO session, the user can simply enter "SOS" and a search expert will respond. The search experts are trained to interpret reference questions, be knowledgeable about available sources of online information, and know how SearchMAESTRO works. DTIC supplies the search experts for the DROLS files, and Telebase Systems, Inc. supplies experts for all other systems.

SearchMAESTRO is currently available to all DTIC government users. These users access SearchMAESTRO both directly and through the DGIS.

SearchMAESTRO is just the first of several end-user interface options we plan to offer STINET users. DTIC has sponsored two conferences in the area of interface technology and maintains a constant alert for new offerings. We are especially interested in expert systems and personal computer-based interfaces.

3. NETWORK FUTURE

Baseline STINET services will be available with the implementation of DGIS, IBIS and SearchMAESTRO. Operational users will be offered access to the network in 1989. With operational status network maintenance will become a major factor. The maintenance required for STINET goes far beyond maintenance requirements associated with hardware and software. It is the intellectual maintenance which will be the challenge.

This challenge will involve tracking changes instituted at remote resources so that appropriate modifications can be initiated on the STINET and network harmony will not be disrupted. STINET maintenance will also involve adding new resources to the network. Information specialists will have to determine what resources should be added to the network and in what priority order. Most importantly, information specialists will be called upon to attest to the validity and reliability of data sources. They will have to ensure that external resources they use, or recommend to their users, are credible. They will have to develop policies concerning how locally generated information will be controlled and validated. This will be critical; the use of incorrect information for decision making and planning could prove disastrous to an individual or to an institution.

As the number of resources available on the network grows, users will want tailored "views" of information personalized to meet their needs. Information "boutiques", organized by subject and interconnected, will form a virtual worldwide, multimedia library. Information specialists will be called upon to organize and oversee these global libraries.

New services must be continually added to the STINET baseline so that the network will keep pace with user needs to the extent that economics and technology allow. One of the most useful services which could be added to this global network would be an automatic translation facility. This would be useful in cases where pertinent information is stored in a language which is not native to the user. Automatic translation would make the information rapidly available in a useful format. The automatic translation service could also be used in conjunction with the "people bases." The translation capability could be used to eliminate the language barrier among users on the network. For example, users could send and receive electronic mail messages in the language of their choice. The technology for this type of service is still in its initial development stages. A limited prototype system is currently in the planning stages and will be evaluated by test users next year.

Optical and video technology will also have a place in the network. One of the questions which must be addressed is when to distribute resources on optical media for local use rather than making the resources available through telecommunications. Distributing information on optical media often has economical advantages when secure communications are required. Optical and video technology bring us into the era of "hypermedia" where we have to redefine our conventional definitions of information systems and the data which they contain.

The challenge of building a global scientific and technical information network will stretch far into the future. Network services will be added and deleted based on changes in technology and user requirements. The unchanging factor in successful development of the

network will be the partnership among end-users, information specialists and network developers. As long as this cooperative relationship continues, the STINET will continue to grow.

Gladys A. Cotter
Defense Applied Information Technology Center
1800 N. Beauregard Street
Alexandria, VA 22311-1784 U.S.A.
(703) 998-4600

**POSSIBILITIES ACTUELLES ET FUTURES DE LA TECHNOLOGIE
INFORMATION TECHNOLOGY: CURRENT AND FUTURE POSSIBILITIES**

Albert Yanez

Centre de Documentation de l'Armement (CEDOCAR), Paris, France

The following is comprised of three items:

The original French language paper by Mr. Yanez,

A translation into English of his paper,

A brief analysis of the needs of a translation capability *vis-a-vis* the DoD Gateway Information System (DGIS), titled *A Look at Some of the Elements of Translation*.

CEDOCAR is the French Ministry of Defense organization comparable in missions and functions to the U.S. Defense Technical Information Center (DTIC). Mr. Yanez holds a position that is second to the Director of CEDOCAR, and interacts with DTIC through mutual participation on NATO/AGARD Technical Information Program panels.

POSSIBILITES ACTUELLES ET FUTURES DE LA TECHNOLOGIE.

Albert Yanez
Centre de Documentation de l'Armement (CEDOCAR). Paris
[Juillet 1988]

L'évolution est très rapide dans le secteur des technologies de l'information, et les possibilités offertes des à présent devraient nous conduire à une réflexion approfondie et à de nouveaux schémas ou modèles de traitement de l'information. Quelques exemples de ces changements peuvent être cités: cf. J. Naisbitt - *Megatrends, ten new directions transforming our lives*, entre autres.

Numerisation:

Aujourd'hui, de plus en plus souvent, les documents sont créés non seulement sur papier mais aussi sous forme numérique lisible par ordinateur, à partir de machines à traitement de texte ou en PAO (publication assistée par ordinateur) (ex; système NPODS de La Navy, ou système Adonis, ou Transdoc...)

Alors que le coût de la saisie de l'information et toutes les contraintes liées à cette étape laborieuse du traitement conduisaient à se limiter à constituer des bases catalographiques ou bibliographiques (titre, auteur, repères divers, court résumé et quelques mots-clés), il apparaît possible et économique aujourd'hui de stocker en ordinateur de texte intégral sur lequel on peut, grâce à des logiciels plus évolués que ceux qui n'opéraient que sur mots-clés, faire des recherches performantes sur tous les codes, caractères, chaînes de caractères, expressions ou phrases significatives du contenu.

Lorsque l'information n'a pas été numérisée dès sa création (cas bien entendu des documents du passé), il est possible de scanner, c'est-à-dire de numériser page par page, image par image, comme le fait par exemple l'office européen des brevets pour soixante millions de pages. Chaque page ou image peut ensuite être compactée, peut être indexée de façon traditionnelle ou de façon automatique grâce à des algorithmes de reconnaissance de caractères associés à des bases conceptuelles, avec un résultat qui peut être supérieur à celui d'une indexation classique.

Parallèlement on travaille beaucoup sur le langage, non seulement la langue écrite mais aussi la langue parlée, et l'on commence à pouvoir reconnaître le discours à partir des enchaînements de phonèmes avec des machines à dicter qui numérisent le discours et opèrent un stockage des communications orales sans passer par un support papier.

En résumé la disponibilité en ordinateur est de plus en plus facile et courante, d'autant que le coût de stockage d'un caractère ou d'une image diminue rapidement. De surcroît les logiciels de recherche évoluent vers une convivialité et une performance "interactive" telles que l'utilisateur final peut y trouver un intérêt, sans recourir à un intermédiaire avec les distortions que cela implique, ni être contraint de passer par le langage artificiel intermédiaire qu'est le thesaurus.

Ceci ne signifie pas que le thesaurus soit devenu un outil inutile ou archaïque, mais simplement qu'il n'est plus l'outil indispensable. Il donnait une représentation a priori du contenu et de l'environnement sémantique de chaque concept représenté par un descripteur, à utiliser pour indexer ou pour rechercher une information. Comme tel il garde une certaine valeur de guide, mais l'effort nécessaire pour l'entretenir doit être pris en compte. À noter qu'on voit se développer des outils de réalisation d'autres formes de thesaurus en opérant à partir du texte lui-même pour faire apparaître a posteriori des relations sémantiques et paradigmatiques s'apparentant à une analyse conceptuelle, à une mise en évidence de la partie la plus

INFORMATION TECHNOLOGY: CURRENT AND FUTURE POSSIBILITIES

Albert Yanez

Centre de Documentation de l'Armement (CEDOCAR), Paris

[July 1988]

The rapidly moving evolution in information technology, and the possibilities it offers now and for the future, ought to have us reflect extensively on the new schemes or models for the handling of information. Some examples of these changes are cited by J. Naisbitt, among others, in his *Megatrends; Ten New Directions Transforming Our Lives*, as follows.

Digitization:

Today more and more, documents are created not only on paper but also in digital machine readable form, including word processors or desktop publishing systems. Examples are: the NPODS system of La Navy, the Adonis system, Transdoc, *et al.*

Although the cost of capturing information, with all its attending constraints at this laborious stage, at one time limited the construction of catalog or bibliographic databases (title, author, diverse markers, abstracts and some terms), it seems possible and economical today to store in the computer integrated text with which one can, thanks to advanced software that functions only against key words, achieve successful searching on all codes, characters, character strings, expressions or significant phrases.

When information has not been digitized at its creation (as is the case with much past documentation), it is possible to scan it, that is, digitize it page by page as is done for example by the European Patent Office for 60 million pages. Each page or image is then compressed, and indexed in the traditional manner, or in an automated manner based on character recognition algorithms associated with conceptual bases with a result that can be superior to classical indexing.

In parallel, there is extensive work being done with language, not only the written language but also spoken language. We are beginning to be able to recognize discourse through phoneme chains with dictation machines that digitize the discourse and function in a storage of spoken communication without passing through a paper phase.

In summary, the availability of automation is becoming more easy and more current, while the cost of character or image storage is decreasing rapidly. What's more, searching software is evolving toward an ease of use and an "interactive" performance to the extent that the end-user can find one's subject without running back to an intermediary with the distortions that all that implies, nor by having to pass through the intermediary artificial language of the thesaurus.

This does not mean that the thesaurus might become an unuseful or archaic tool, but simply that it is no longer an indispensable tool. It gave an *a priori* representation of contents and of the semantic environment of each concept represented by a descriptor, to be used for indexing or for searching. As such it retained a certain guide value, but the effort needed to maintain it should be taken into account. Tools are being developed for creating other thesaurus forms based on the text itself to have *a posteriori* the semantic and paradigmatic relations coming from a conceptual analysis, which places in evidence the most significant information content, of a sort that what results can have a new interest, notably for inquiry, to approach the techniques used in bibliometrics, infometrics (*infometrie*), and scientometrics (*scientometrie*) . . . for example one can see appear a nonfortuitous link between carbon film and furtivity that a thesaurus

significative du contenu informationnel, de sorte que ce qui en résulte peut avoir un intérêt nouveau, notamment pour le renseignement, à rapprocher des techniques utilisées en bibliométrie, infométrie, ou scientométrie... par exemple on pourra voir apparaître un lien non fortuit entre film de carbone et furtivité qu'un thesaurus établi a priori n'aurait pas identifié...

Microinformatique:

Les progrès associés à la microinformatique font apparaître aussi une autre tendance très nette: alors que l'on s'efforçait de regrouper sur un même site informatique puissant et central les banques de données textuelles et même factuelles, pour aboutir à la notion de serveur, on découvre aujourd'hui des avantages à s'orienter vers une informatique répartie, ou en réseau. En effet un microordinateur peut avoir une capacité suffisante, avec des disques durs ou compacts, pour recevoir une ou plusieurs banques de données entretenues et utilisées par une communauté particulière qui n'est plus tributaire d'un système centralisé plus contraignant. Le même microordinateur permet l'interrogation de plusieurs serveurs et l'information regroupée par téléchargement peut être utilisée localement sur un site protégé, si elle est mêlée alors à de l'information protégée. Ou bien on peut juger préférable de ne pas passer du tout par des lignes téléphoniques et d'échanger des supports magnétiques (disques ou bandes). Autrement dit des décisions sur le traitement de l'information protégée sont très dépendantes des aspects technologiques.

Reseaux:

Une autre évolution très rapide est à prendre en compte, et elle accélérera la disponibilité de l'information contenant du texte et/ou des images. La prolifération des réseaux publics et privés et des services de messagerie ou courrier électronique utilisant de larges bandes et des technologies très fiables permet de transporter du texte intégral et de procéder à des transferts électroniques de documents la ou pour l'instant on ne transférait que de courts messages. À titre d'exemple le réseau numérique à intégration de services (RNIS) commence à être opérationnel et sera d'un emploi généralisé en 92. La transmission à 64 kbits/seconde laisse entrevoir de nouveaux comportements, par exemple l'utilisation de serveurs de traduction assistée (Systran, Alps, Wiedner... et Eurotra), le texte source étant transmis et la traduction renvoyée immédiatement. Une autre conséquence de cette capacité accrue des réseaux de télécommunications est déjà l'apparition de passerelles (gateways) entre serveurs qui s'appuient sur des réseaux de courrier électronique (comme Infotap et Geomail), et des possibilités de transferts rapides entre collections archivées sur disques et juke-boxes, avec des protections par algorithmes complexes, comme c'est le cas de SARDE, commercialisée à partir de recherches du CNET.

Bien d'autres exemples d'évolutions allant dans le même sens pourraient être cités. En conclusion il semble que l'évolution technologique dans le secteur des technologies de l'information devrait être prise en compte dans une réflexion de type systémique, de manière à essayer d'intégrer de nombreuses données conditionnant le processus de décision ou de définition d'une politique générale ou d'une stratégie de l'information de L'OTAN.

established *a priori* would not be able to identify.

Microcomputing:

The progress associated with microcomputing has also given rise to another very clear tendency: where before one tried to group textual and factual databases in a central location with the idea of having centralized processing, today one discovers the advantages of leaning toward distributed processing or networking. In effect, a microcomputer can have sufficient capacity, with hard or floppy disks, to access one or several databases maintained and used by a particular community that is not dependent on a more constraining centralized system. The same microcomputer allows interrogation of several hosts, and the information aggregated through telecommunication can be processed at the local site, even in being combined with information already stored in the micro. Or one can decide not to make use of telecommunication access but instead exchange information electronically stored on disks or tapes. It can be said that the decisions for processing local information are very dependent on the developing technology.

Networks:

Another rapidly evolving technology is to be taken into account, and which is accelerating the availability of information in text and/or images. The proliferation of public and private networks and electronic message or mail services, using wide bands and highly reliable technologies, is allowing the transfer of integrated text; this permits the electronic transfer of those documents whereas at one time one sent only short messages. An example is the *services integration digital network* (reseau numerique a integration de services (RNIS)) which will be operational for general use in 1992. Its transmission at 64 kbits per second lets us have a glimpse of new capacities, e.g., the use of assisted translation hosts (Systran, Alps, Wiedner...and Eurotra), the text source being transmitted and the translation sent on immediately. Another consequence of this capability because of telecommunications networks is the appearance of gateways (*passerelles*) among hosts that rely on the electronic mail networks (such as Infotap and Geomail), and the possibilities of rapid transfers between collections archived on disks and juke-boxes, protected by complex algorithms, as in the case of SARDE, which is a commercialization of the research done by CNET.

Many other examples of technology evolutions taking place in the same manner could be cited. In conclusion it seems that the technology evolutions going on in the information technology sector ought to be reflected on relative to the system development, in the sense of trying to integrate the numerous data present that effect the decision process, or the definition of a general policy, or a NATO information strategy.

(19 Aug 88)

A LOOK AT SOME OF THE ELEMENTS OF TRANSLATION
Preparatory to Accessing Global Information

Allan Kuhn
Hypermedia Laboratory

An intent of future DoD Gateway Information System (DGIS) development is incorporating human language translating capabilities in the global accessing and transfer of information. We envision a system that provides for both passive and interactive machine translation and transmission. The areas that we see immediately useful are:

- Translation of foreign language information retrieved from foreign systems.
- In-process translation into English of messages received on the electronic mail from foreign language speaking people, with reciprocal translation from English.
- Communicating interactively with foreign language databases, with interactive translation between English and the language of that database. An extension would be interactive translation between a selected language and the language of the database.
- Interactive translated communication with foreign language speaking people while making use of the DGIS capability to "talk" interactively via the keyboard. An extension would be translation of or to voice-generated data.

In light of today's technologies, DGIS may want to make use of both external and internal translating capabilities. We may want to incorporate human language translation interfaces on DGIS, we may also want to establish a link with a translation ordering service, and we may want to integrate into a network that includes machine translation nodes.

Except for the first translation area of interest above, the other areas deal entirely with machine-controlled translation. In the preceding paper by Mr. Albert Yanez, he makes note, for example, of "assisted translation" servers in Europe, such as Systran, Alps, Wiedner, and Eurotra. Text may be directed to include such a server in its transmission path, and the text is translated and sent on to its destination.

Although the elements of human and machine translation are well established, we did a translation of Mr. Yanez' paper as an exploratory exercise to get a grasp on the needs for translating, and consequently to establish at least a small knowledge base of those needs, as applied in the area of conceptual needs for machine translation. We did not make use of machine translation capability in this exercise. What we wanted to learn were concepts involved in translation, in order to be able to discuss how the technology could be initiated for DGIS and what would it take to have that technology.

We also wanted to see what machine technologies available at the DAITC could be used in doing a translation. So, to facilitate the human translation of Mr. Yanez' paper we tried to incorporate several technologies. The three basic technologies were:

- o scanning the original French language paper into bitmapped graphics format,
- o conversion to ASCII digital format for first level French language flagged character editing,
- o and as a result of the digital format further handling the text with computer text editing capabilities during the translation to English.

Computer text editing includes highly useful features such as global edits, needed to make sure that all terminology is correctly correlated in the translation.

The systems involved were:

- o Palentir OCR System - for putting the text into graphics format;
- o SUN 3/260 - our bitmap system for handling the graphics format;
- o DOCPRO - Software for conversion to digital format and editing;
- o DoD Gateway Information System (DGIS) - UNIX text editing.

Within the narrow scope that this translating exercise was done, the following elements surfaced. Some of these elements are very obvious.

ELEMENTS

Information destined for translation has two basic sources: paper and electronic input. In a sense, because of today's technologies, the electronic input is an extension of the paper source. Because of that, the translation elements of paper source translations are also applicable to electronic handling of items directed toward translation.

A. Machine Technologies

In all machine handling of human language text, the character set of that language must be present. Although this need is obvious, this requirement begins at the optical reader for machine paper input, is retained throughout the machine handling of the language text, right to the point of translation. An optimal machine input of a native language piece cannot take place without the language character sets; otherwise extensive human knowledgeable intervention is required.

In optical reading, the machine reader must be able to recognize fonts that are characteristic or cultural to the language. Cultural font print sources range from typewriters, through electrical printers, to mechanical printing machines. If font recognition is erratic, human editing intervention is not only intensive but also improvisational. An example of improvising is doing without all the French language accent marks, as shown by the French section of Mr. Yanez' paper. An example of a cultural font is:

texte source etant transmis et la traduction renvoyée
immédiatement. Une autre conséquence de cette capacité accrue des
réseaux de télécommunications est déjà l'apparition de
passerelles (gateways) entre serveurs qui s'appuient sur les
réseaux de courrier électronique (comme Infotap et Geomail), et
des possibilités de transferts rapides entre collections
archivées sur disques et juke-boxes, avec des protections par
algorithmes complexes, comme c'est le cas de SARDE, commercialisé
à partir de recherches du CNET.

Example of a cultural French typewriter font.

Additionally, reliable discernment of each print character by the optical reader is highly desirable. Without it, human editing intervention is again extensive.

Associated features in optical scanning concern the conversion to the ASCII format. The software programs used copied the bitmap format correlative to line length, spacing between sentences, and misread characters. The misread characters, caused by the lack of the native language character set, were inevitable. But the ASCII duplication of the bitmap format in lengths and spacing also required extensive human editing of the French text.

texte source etant transmis et la traduction renvoyee
imm^{me}ediatement. Une autre consequence de cette capacite accrue des
reseaux de telecommunications est deja l'apparition de
passerelles (gateways) entre serveurs qui s'appuient sur des
reseaux de courrier electronique (comme Infotap et Geomai^{it}), et
des possibilites de transferts rapides entre collections
archivees sur disques et juke-boxes, avec des protections par
algorithmes complexes, comme c'est le cas de SARDE, commercialise
a partir de recherches du CNET.

Example of conversion to ASCII, duplicating graphics format (with some second level edits).

The first level editing of the French text was done through the DOCPRO flagging feature, and when that was completed, the text was transmitted to the DGIS for further editing (second level), and human creation of the translation. A projected technology application that came out of the exercise, however, would be making use of the windowing capability of the SUN (or any other bitmap system) for simultaneous views of the native language and the translation.

B. The Human Translator

In human translation or human monitored translation, there are two factors involved. One is the style of the author, the other is the style of the translator. Obviously, in true translating, the translator must repress his own sense of style. On the other hand, however, the translator must discern the style of the author in order to carry that style into the translation. But the translator also finds himself making judgements about the author's style as to how much to carry into the translation the author's style characteristics. An example is overly lengthy sentences. A dilemma that the translator faces is whether to perform as true a translation as possible, to represent the cultural style of the native language, or to break the lengthiness for the cultural English language representation of the translation for the purpose of achieving a greater clarity in the thoughts communicated. To modify the text in the translation becomes a very subjective issue.

Another element in human translation is discerning the audience of the author and that of the translator. Although one would expect, for example, that the audiences for a technical piece would be the same, cultural differences can also create contextual differences in perceptions of the piece. Again, a dilemma of the translator is making a true translation *vis-a-vis* a contextual translation in terms of the culture.

Topical terminology also presents a contextual problem. The translator must not only be familiar with the language and its culture, but also with the technology in both languages. The technology terminology of the translation language must be identified and correlated with the correct technology terminology of the native language. If a translator is doing a piece on avionics engineering from German to English, he must not only know both languages but also the technical language of avionics engineering in both languages.

The translator, therefore, must face these issues and make judicious decisions in achieving clarity for the translation in the eyes of the translation's audience. Included in this process is assuring that the author's

intent is carried across in the translation to the translation's audience.

HOW THIS RELATES TO MACHINE TRANSLATION

A machine detects neither style nor context. The receiver of a machine-controlled translated item must realize, therefore, that the translated piece at hand is likely to be out-of-context in parts, and that the context is difficult to discern without the original piece in hand and knowing the language. If the receiver sees strange turns of terminology, for example, he must either attempt a rational deduction as to the actual terminology, or be at a loss as to the true terminology intended.

For a formal translation, the human translator must be at both the in-side and the out-side of the translation, for it is the translator that exercises the human cognitive characteristics of:

- o understanding the native language context of the paper,
- o discerning the intents of the author,
- o verifying the correct correlations in the technical terminologies of the two languages,
- o verifying the correctness of the context of the translation.

In the case of the DGIS translation interface goals mentioned at the beginning of this paper, the absence of a human translator means that the above human characteristics cannot be employed in machine translation. Furthermore, the machine translation cannot be verified against the original piece at the front end of the translation process, including an electronic item. Consequently, some lack of translation reliability must be both expected and accepted. In the case of an electronic mail message, for example, a criterion might be a translated message that is reasonably understandable. A higher expectation of reliability would be wanted in translation of information from a foreign language database.

In spite of the elements above effecting machine-controlled translating, I would recommend that machine translating capabilities be incorporated into DGIS. The goal of being able to communicate globally would then be established for DGIS communications. This exercise was also performed within the context that machine translation is one of the outgrowths of the development of artificial intelligence. I would further suggest, therefore, that refinements would follow; refinements supported by developments making use of artificial intelligence tools to established information-in-context, rather than simply streams of data.

ACKNOWLEDGEMENT

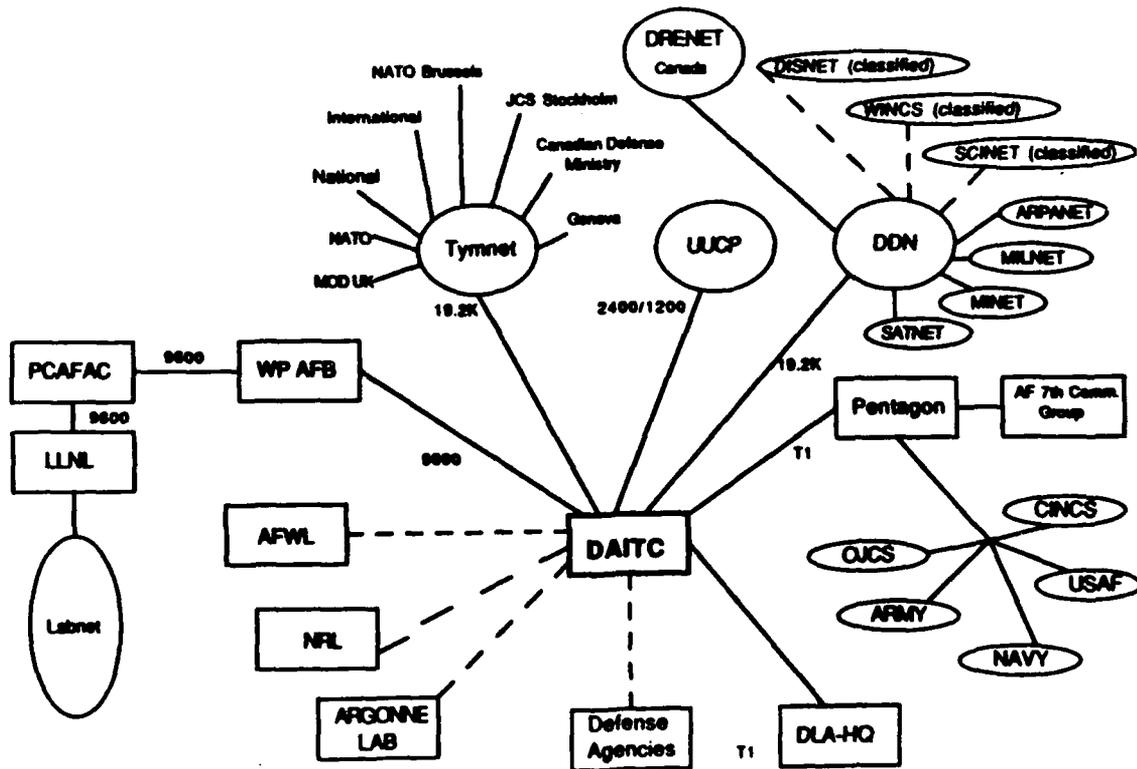
The machine optical input and edit preparation described in this paper were done with the assistance of the DAITC Text Search Laboratory, under Ms. Karen Kaye and Mr. Joe Morton.

**NETWORKING AND INTEROPERABILITY LABORATORY
DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER**

*Judy F. Hunter
Director
Networking and Interoperability Laboratory*

[November 1988]

The section following relates the DAITC networking areas of interest being established as a global means of internal and external information transfer, and is reproduced for inclusion in this report.



NETWORKING AND INTEROPERABILITY LABORATORY

Laboratory Mission

The mission of the DAITC Interoperability Laboratory is to provide an environment for the assessment of contemporary trends in networking and data communications in order to furnish plausible solutions to demands for direct and transparent access to a wide variety of geographically distributed resources. The lab will serve as a central source of information with a pool of expertise in communications as well as a repository of interoperability solutions that can be applied generically to problems within the DoD community. Basically, the premise is that individual DoD communities are working to find solutions for the same types of communications problems, in isolation, which is resulting in a duplication of effort and funding among DoD agencies. By establishing a centrally located and accessible laboratory, DoD can save the costs of expensive computer prototypes and personnel by sharing the knowledge among those agencies with similar requirements. Even if solutions are not applicable to a specific situation, parts of that solution might be applied successfully.

Laboratory Goals

1. **Provide a DAITC communication infrastructure to include the hardware, software, application of communications standards, and telecommunications capabilities.**
 - o **Coordinate DAITC networking issues.**
 - o **Oversee continued maintenance and development of the existing LAN and WAN**
2. **Provide generic solutions to DoD interoperability problems, including multi-level and secure networking.**
 - o **Remain cognizant of DoD interoperability requirements.**
 - o **Provide DAITC sponsors with a pool of expertise and a repository of information about interoperability issues.**
 - o **Record the results of research and the development of prototypes for dissemination to DoD agencies.**
3. **Establish DAITC presence on communications standards committees.**
 - o **Assess contemporary trends in communications technology and standards in order to apply these trends to current problems within the DoD community and in planning for the future.**

- o Remain involved with those agencies involved in setting policies for networking and telecommunications in the Federal government, including those evaluating the Government Open Systems Interconnect Protocol (GOSIP) and the National Bureau of Standards (NBS) forums for OSI and the Transport Control Protocol/Internet Protocol (TCP/IP) evaluations.

Hardware and Software Available and Under Evaluation

The DAITC has an established Ethernet-based local area network (LAN) that utilizes a variety of state-of-the-art communications technologies. The LAN is running the Transport Control Protocol/Internet Protocol (TCP/IP) and Xerox Network Service (XNS) protocols. The network is interconnected in three hierarchies as follows:

1. The **Frontend Segment** of the network handles access for all incoming traffic to the network. The LAN is accessible using direct dial-in methods, such as modems or Tymnet, and by multiplexor and hardwired connections.
2. The **Backbone Segment** consists of the host computers as well as the T1 links that have been established between the DAITC LAN and the DTIC and Pentagon LANs. The dial-out and outgoing Tymnet lines are also located on this segment of the LAN. The host computers are connected to the Ethernet Ethernet controllers including a VAX 11/780, a Pyramid 98x, a Gould 6050, and an Elxsi 6400. These hosts are loaded to perform specialized functions for prototypes under development currently. The Pyramid 98x currently serves as the gateway for the Defense Data Network (DDN) node at the DAITC.
3. The **Development Segment** provides connectivity to the Backbone and Frontend portions of the LAN for the research and development labs at the DAITC. Since this portion of the network has been set aside for those projects under development, it is possible to isolate this part of the network entirely or specific labs from the Backbone Segment if necessary to prevent impact on the main LAN.

These segments are connected together with Bridge Communications hardware to control access to local resources.

Current Projects

Interoperability issues under evaluation currently include:

- o DTIC Macintosh Network Evaluation to incorporate Macintosh microcomputers into the DTIC LAN so that they can communicate successfully and usefully with other IBM-compatible microcomputers, with host computers, including the DLA DTIC DMINS machines as well as the host computers available at the DAITC.

- o DLANET analyses and the evaluation of possible protocol converters to provide a link at the DAITC to the DLANET. This project will provide DINET, located at the Skyline building, with access to this network initially and will pave the way for future DoD client access.
- o Secure LAN prototype has begun analysis and design phase.
- o Design, develop, and test the concept and deliver an interface allowing selected FORDTIS users who are connect to the FORDTIS DEC VAX VMS cluster system to have direct access to the classified DROLS operating on the DTIC Sperry system. The interface will include a JBM Electronics synchronous/asynchronous converter to be located in the FORDTIS central site. This converter will enable FORDTIS DEC VT100 terminals to communicate with DTIC's Sperry hardware. The communication link will be a 9600 baud dedicated circuit terminating at the DROLS and FORDTIS central sites.

Completed Projects

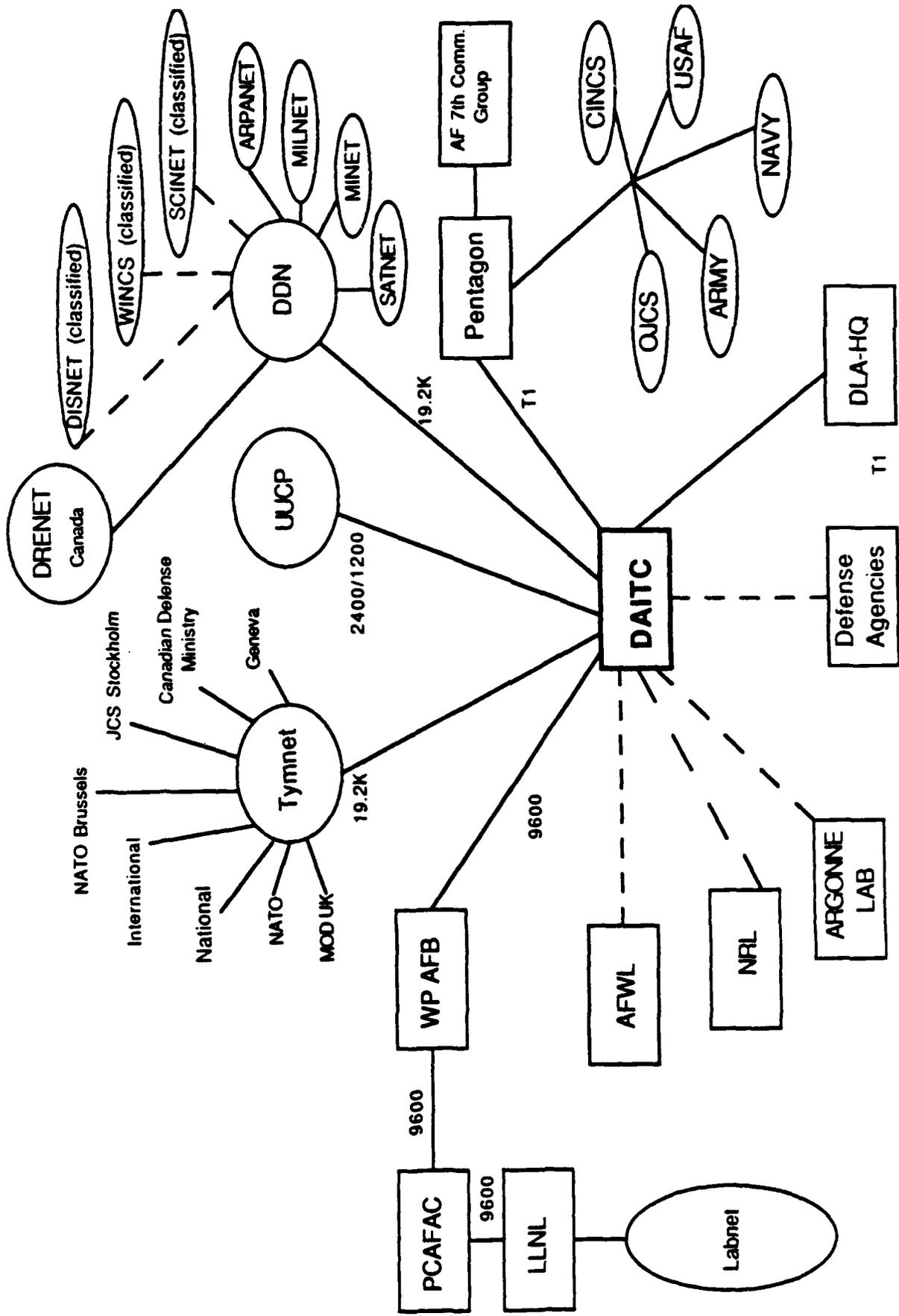
- o DLA E-Mail 8.1 Modification for use at the 1988 DLA Commanders' Conference on April 25-29. The lab completed a project to make the E-Mail 8.1 electronic mail software compatible with the Quadratron Q-Mail software so that both packages can run on the same DMINS machine.
- o Completed the testing and evaluation of a Unix-to-Unix-Copy connection between the VAX 11/780 at the DAITC and the DTIC Gould 90xx machine (DMINS) at DTIC. The operational link provides connectivity that allows users on each machine to transfer electronic mail and files between them.
- o The establishment of a T1 link between the DAITC and the Pentagon has resulted in a requirement to evaluate the problems associated with running multiple XNS protocol based networks and database servers over one Wide Area Network (WAN).
- o DLA UUCP-DDN connection to provide DLA DMINS users with access to the DDN node at the DAITC in order to send and receive electronic mail.

Future Projects

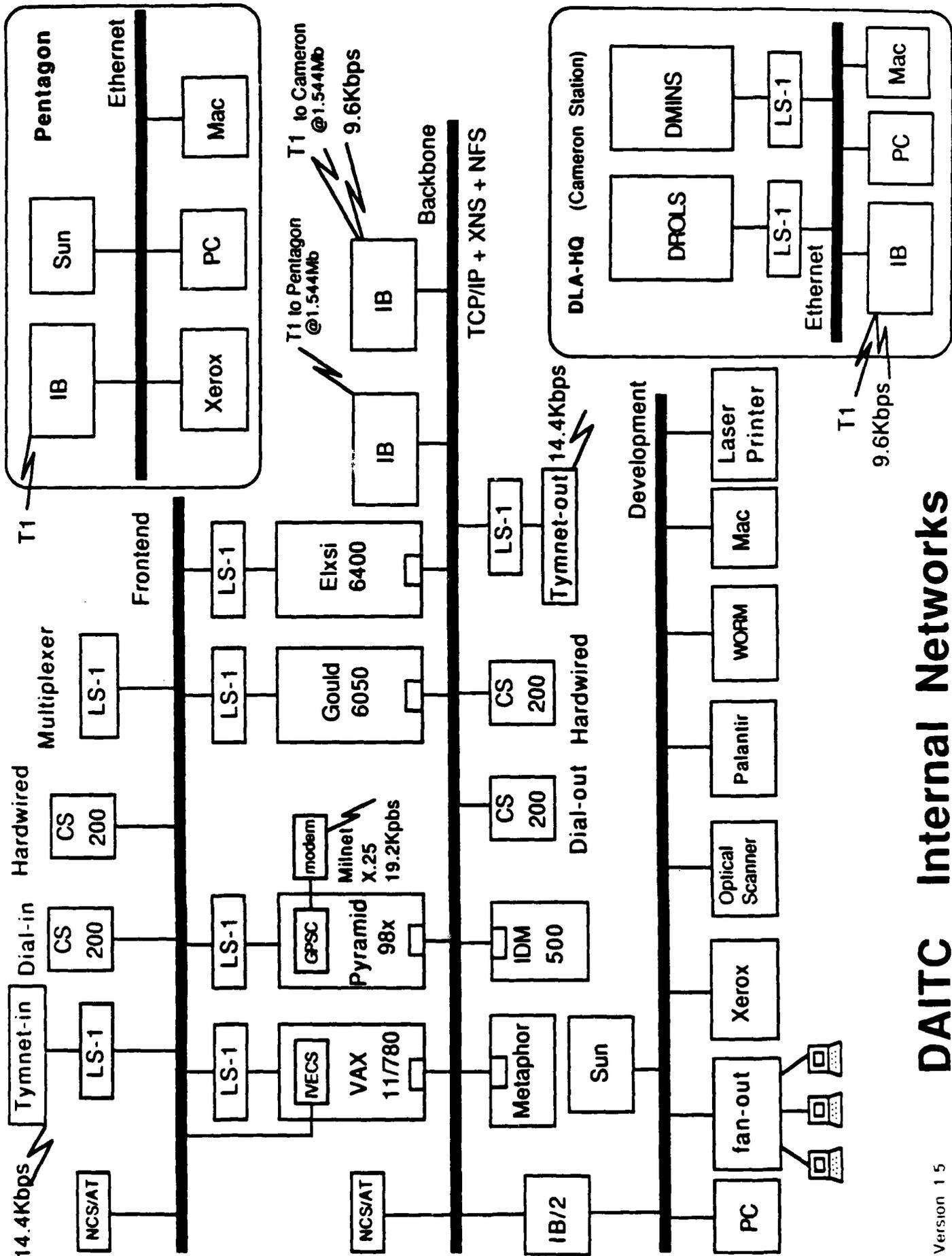
Some of the projects under consideration for future development include:

- o Secure multi-level operating systems and connectivity
- o Interoperability between homogeneous database machines, such as Wang, Sun, Xerox, and Dec.
- o Electronic mail compatibility issues between various types of mail standards, including the evaluation of X.400.
- o Expert System for Network Management

For further information please contact: Judy Hunter (703) 998-3450

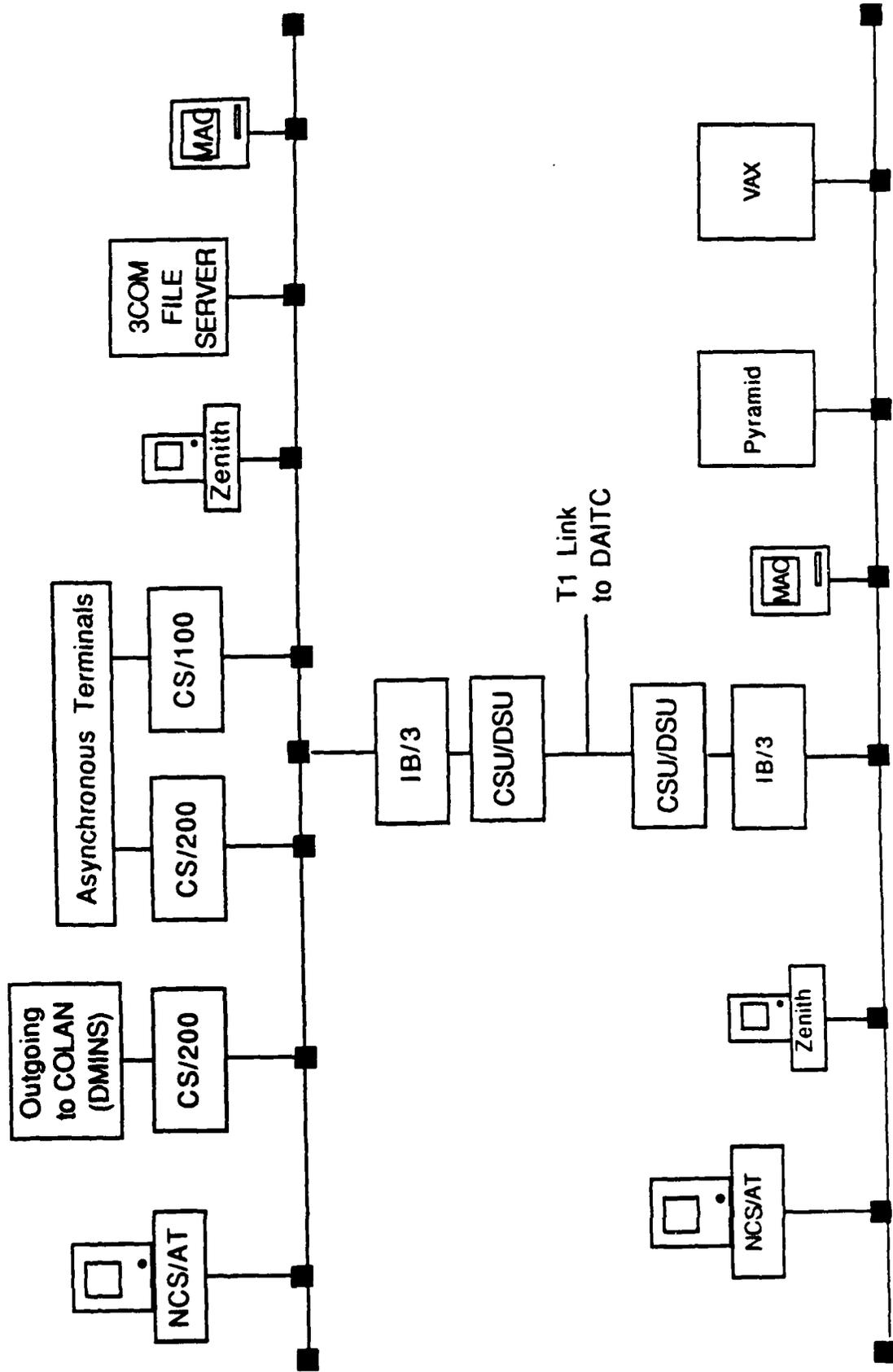


DAITC External Networks

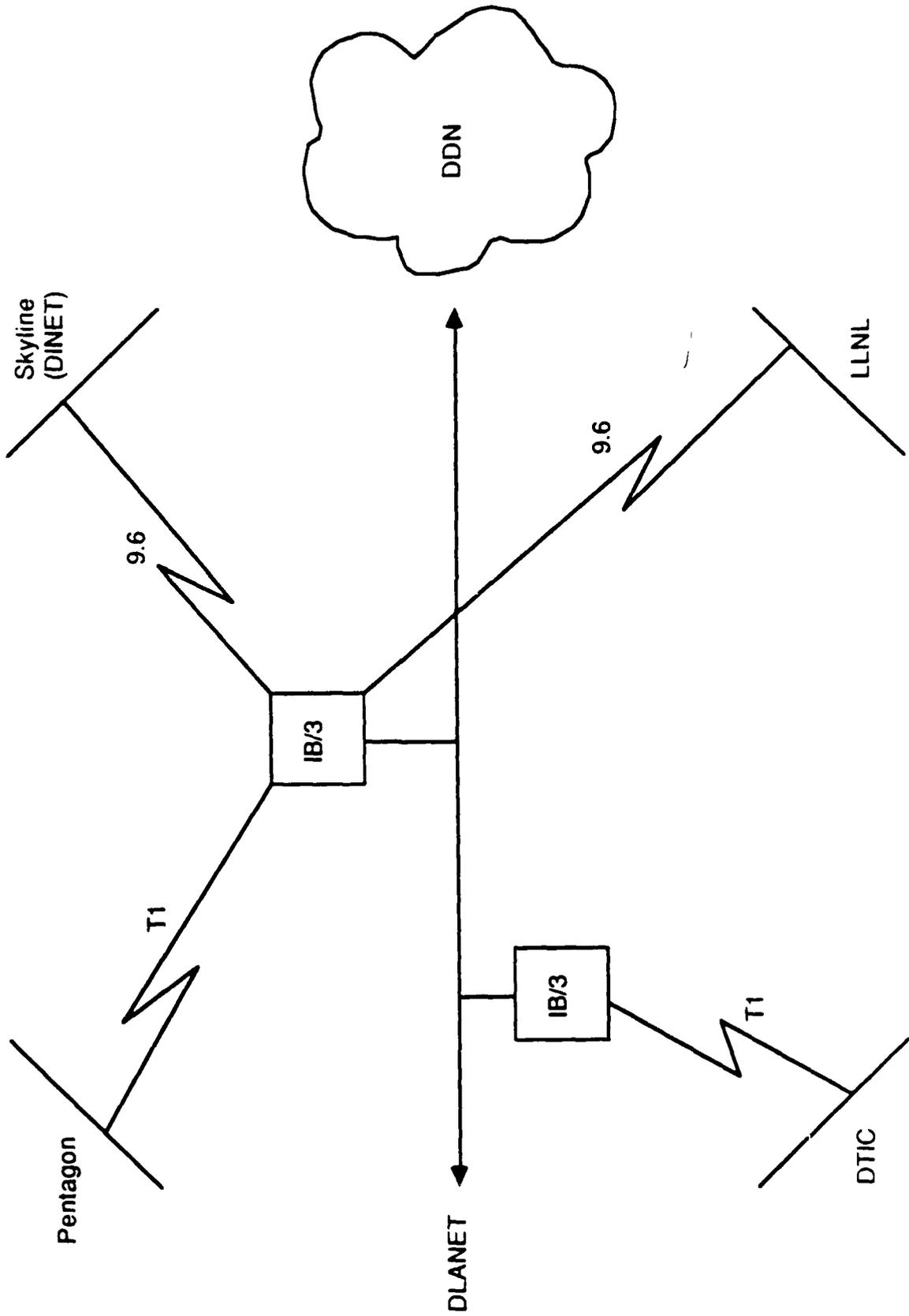


DAITC Internal Networks

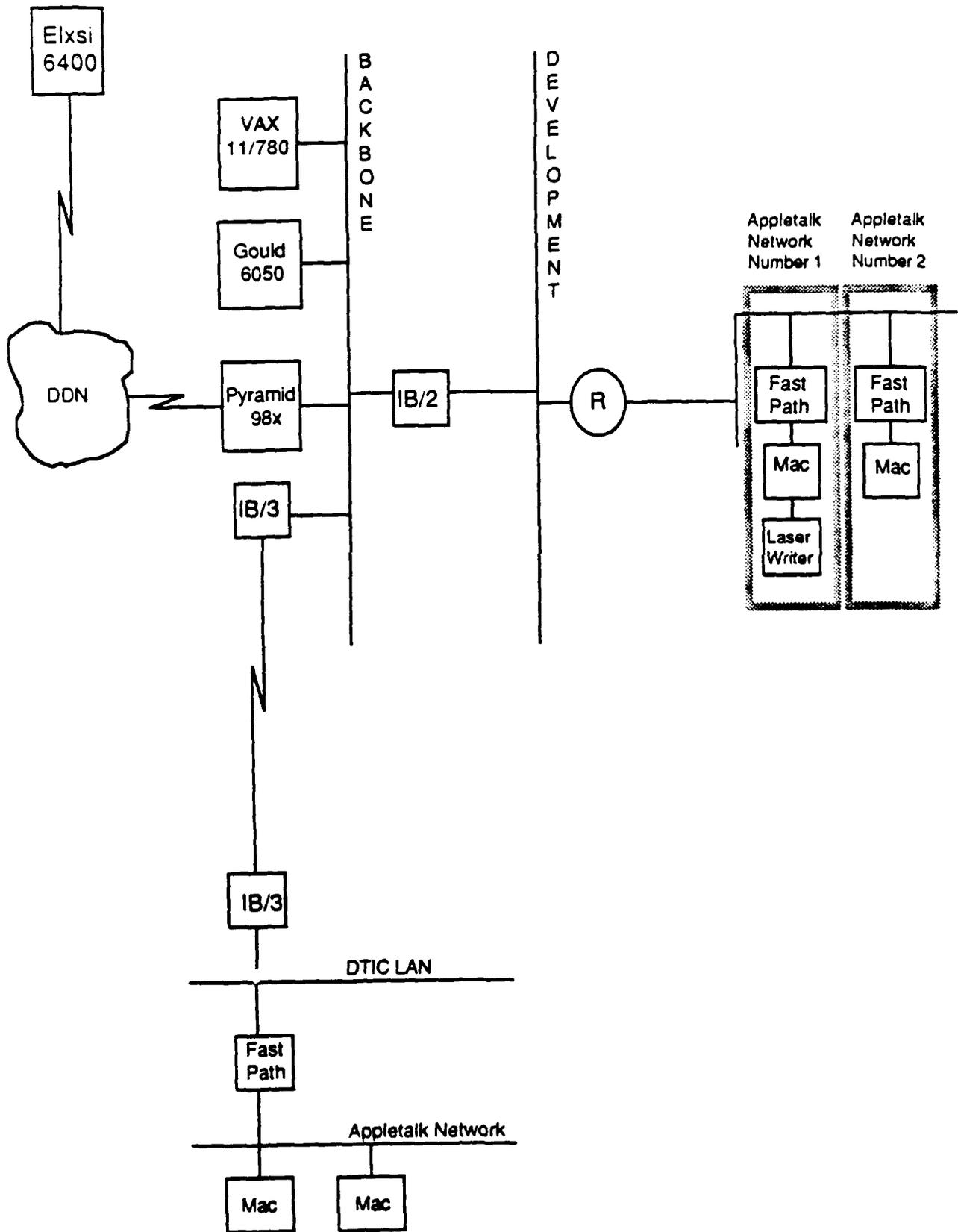
Internet Bridge



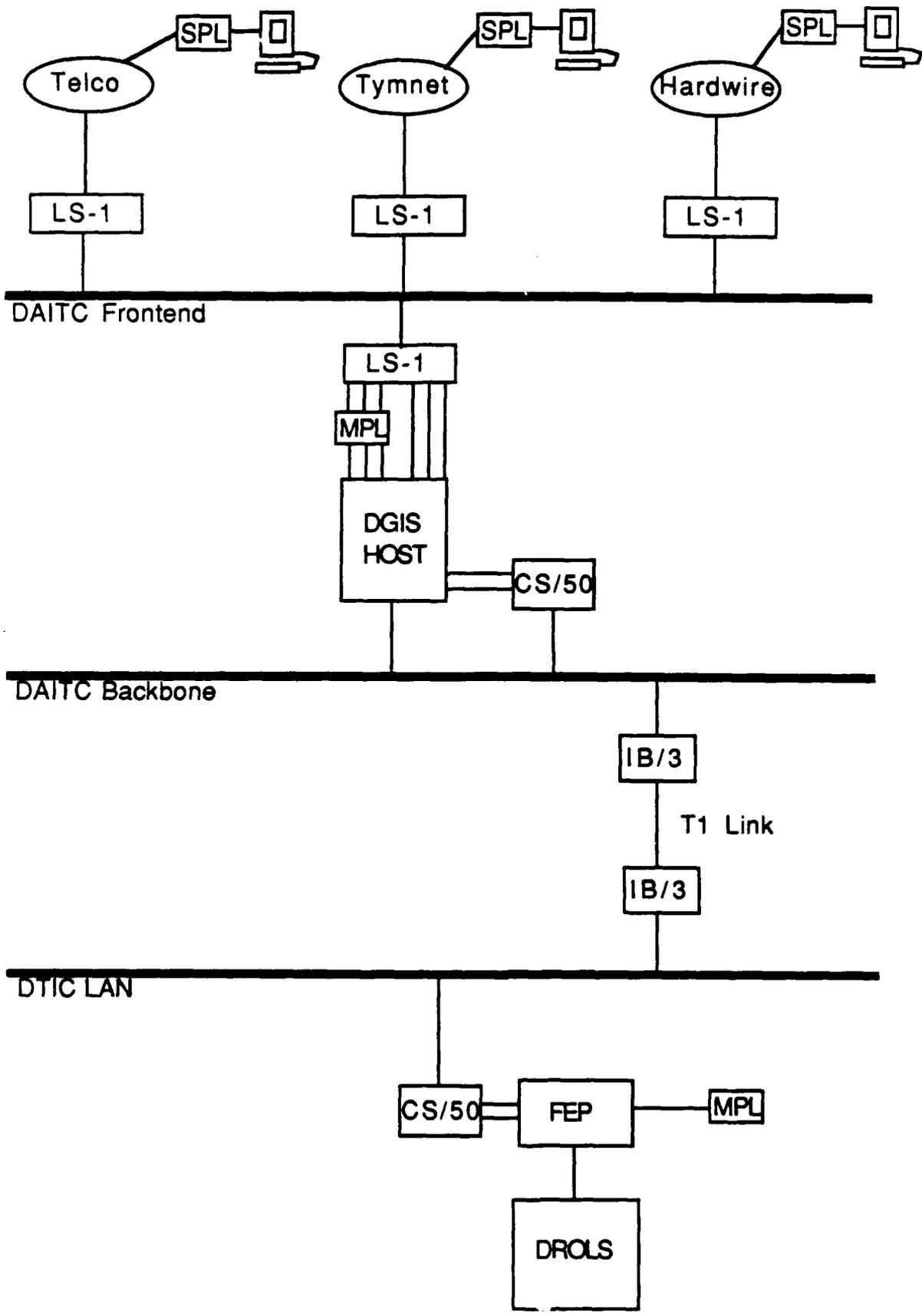
DLANET



Macintosh / FastPath Project



DGIS LEAD IMPLEMENTATION PROPOSAL



DAITC NETWORK EXPERT SYSTEM

*Brenda A. Edleson
Hypermedia Lab.*

Date: Fri Jul 22 14:44:21 1988 From: bedleson (Brenda Edleson) Subject: DAITC Network Expert System

As a result of the Hypermedia Lab meeting of 21 July, 1988, a concept for a DAITC Network Expert System (DANES, for now) was proposed.

The objective of the DANES is to develop a joint DAITC Labs application utilizing Artificial Intelligence and Hypermedia techniques to provide a higher level of productivity and technology within the DAITC. The targeted users of the DANES are the DGIS/DAITC systems staff. The primary uses of the DANES would be network management, planning, and hypothetical analysis.

The DANES would be developed on the Symbolics 3650 using the KEE (Knowledge Engineering Environment) expert system development tool in the Hypermedia Lab. The Networking and Interoperability Lab and the DoD Gateway Information System (DGIS) Lab would provide the expert knowledge needed for the expert system knowledge base.

At the root of the DANES would be a graphical representation of the DAITC network with an associated object oriented knowledge base. Some of the information in this knowledge base would include, for example:

- o System terminal characteristics (type, baud rate, communications server, *et al.*),
- o Linkage information,
- o Characteristics of systems on the network such as addresses and system types.

On top of this, knowledge base rules would be added to capture the relationship of the components and to use this data to do management, planning, and analysis functions. Examples of possible functions are:

- o What terminals are connected to comm server 1?
- o What are all the terminal lines connected to the network and what are their physical locations?
- o What would happen to the load on the system if we pulled comm server 2 and moved its lines to other available ports?

Besides the general systems type functions, the DANES would provide a full scaled documentation capability. KEE provides a very good environment to interweave text and graphics throughout the DANES in a nonlinear way, resulting in a Hypermedia system. An example would be when a user is not familiar with the characteristics of a particular system on the network, the user could click on that system on the network map and a description of that system would appear in a new window. Then the user could traverse through highlighted objects to get more detailed information on particular subjects.

All the aforementioned ideas would be implemented in the initial phase of development. In addition to the above core system, the desire was expressed to expand the proposed system to include several added capabilities including:

- o Migration to different workstations (i.e. Sun, MAC II, IBM PC 386),
- o Taking action as a result of data gathered by network monitoring devices,
- o Implementation on different networks,
- o Internal nam generator.

The proposed addition above of taking action on data gathered from device monitoring, would be a real time expert advisory system which would obtain data from online monitoring devices connected to the network on a continuous basis. In addition to collecting data, the expert system would advise the user as to actions that should be taken to avoid problems and alert the user when an alarm condition exists. It would continue to follow the progress of the situation and update or retract the advice given.

In summary, the proposed DAITC Network Expert System would centralize the expertise of the network systems staff into one application capable of providing the naive and the expert user the ability to look at and analyze the DAITC network in a multitude of different ways. With the DANES capabilities, the DAITC would be a truly advanced network monitoring and management facility, utilizing a combination of communications, Artificial Intelligence, and Hypermedia expertise.

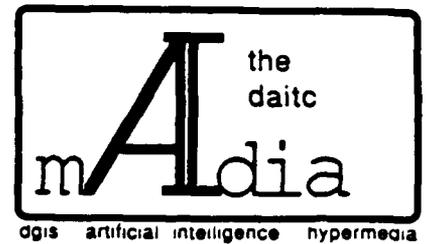
STAFF COMMENT:

Date: Thu Aug 4 15:25:43 1988 From: sstovall@daitc.daitc.mil (Steven Stovall)

I believe that DEC developed an expert system for system and/or network configuration that could provide a point of departure. It could be an interesting project. The DAITC network configuration is becoming quite complex: we have multiple ethernets separated by, e.g., T1 links. There are multiple address spaces in the collection. The collection is connected to the internet via DDN. We query and respond to the internet distributed databases for routing and name service. The complexity curve will be rising steeply for the foreseeable future.

Date: Fri Jul 22 14:53:31 1988 From: jhunter (Judy Hunter)

I think either the Army or the Navy has a DAIN program (Defense Army Intelligence Network maybe) or something like that. If we start throwing the sound DANE around, it might be misconstrued. I assume that we will market (not for sale) this product down the line.



X-WINDOWS

Literally, a window to the world

X-Windows is a software program for distributed bitmap system information handling. It is a way to network graphics programs and systems. In development since 1984 at the Massachusetts Institute of Technology, it is now getting a lot of notice in the technical press.

X-Windows provides high-performance, high-level, device-independent graphics. A hierarchy of resizable, overlapping windows allows a wide variety of application and user interfaces to be built easily. Network-transparent access to the display provides an important degree of functional separation, without significantly affecting performance, that is crucial to building applications for a distributed environment. To a reasonable extent, desktop management can be custom tailored to individual environments, without modifying the base system and typically without affecting applications. (Robert W. Shiefler, MIT Lab. for Computer Science, et al., "The X-Window System," Jul/Oct 86; paper prepared for Transactions on Graphics #63, Special Issue on User Interface Software, copyright ACM. Non-commercial copying granted by ACM.)

X-Windows* was developed at MIT as part of Project Athena, a three-way partnership of MIT, DEC, and IBM. 19 Major computer companies are interested in the system, and it is a *de facto* standard submitted to ANSI. Amazingly, MIT is making this development an open, public-domain standard that provides a single programming and user interface to the networked, bit-mapped graphics that make up windowing systems. X plays to the requirement for portability in a heterogeneous environment.

X contains two basic software components: Xlib and X Server. Xlib resides with an application, while X Server resides in the user's display device. Xlib passes screen display information from a host application to X Server, and passes user-input information from X Server back to the application host. X hides communications from the application and user display, maintaining open windows on a single user workstation to applications on a variety of remote, heterogeneous computers. X can use any error-corrected duplex byte stream from asynchronous communications, allowing use of a tree structure and thereby permitting an infinite nesting of windows.

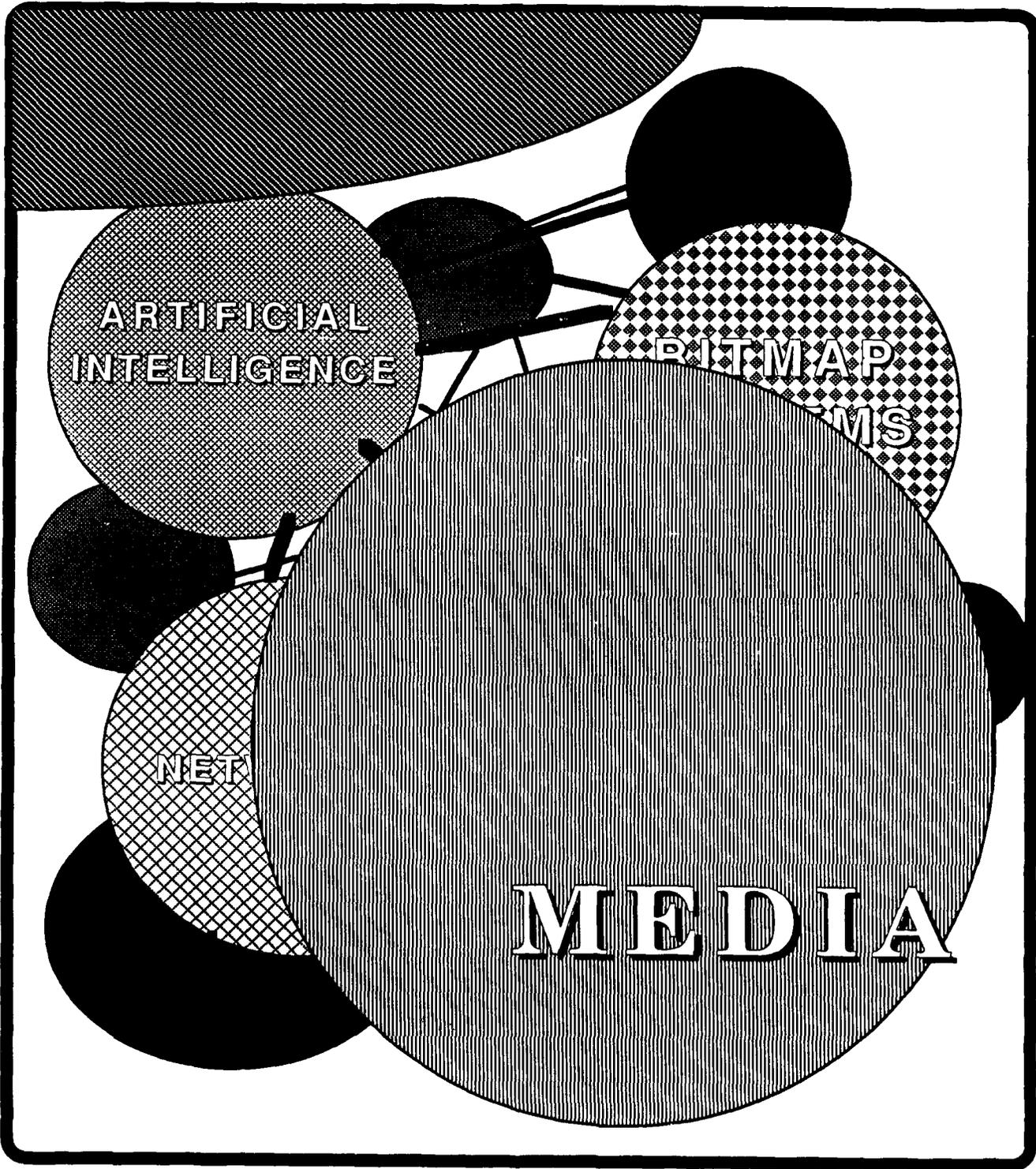
Some companies, in supporting this development in the form of a consortium, are beginning to come out with their own versions of this system-independent heterogeneous processing software. An example is DECwindows, marketed as a network-transparent universal interface, making available applications that users need not even run on their own systems.

A spinoff is the development of X terminals. Some are low-cost bitmap terminals comparable in cost to ordinary terminals. They provide the Xlib capability to display and communicate with the X Server. They are now offered by vendors ranging from the low-cost Acer Zebra 100 to X Window System implementation of Common X interface for HP workstations *et al.*

The Hypermedia Lab's interest in X-Windows is in incorporating it as a component into DGIS hypermedia processes. X is critical in establishing a standardized hypermedia-based STINET for the contextual handling of information, as a support adjunct to the DGIS mission. DoD bitmap system users would have a standard, resource-sharing hypermedia-based DGIS/STINET at their disposal through X-Windows.

** X-Window descriptive information following is taken verbatim from "X Primer: Closing the User Knowledge Gap," Edith D. Myers, UNIX Today!, 14 Nov 88. Additional source: "The X Window System," Dick Pourtain, Byte, January 1 89.*

The prime goal of the Hypermedia Lab. is to create AI-based programs that make the human-machine interface more human in the machine's responses.
DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER, ALEXANDRIA, VIRGINIA



ARTIFICIAL
INTELLIGENCE

BITMAP
SYSTEMS

NET

MEDIA

DAITC HYPERMEDIA LABORATORY

HYPERMEDIA IN DOD Where The Medium Is Definitely The Message

Brenda Edleson
Hypermedia Laboratory

November 1988

Articles appear frequently in both the popular and technical press about the incorporation of innovative information technologies in Defense activities. An example of one such article, appearing in *AI Magazine*, Fall 1988, is "What AI Can Do for Battle Management." (1) This article talks about current developments that are generally known, along the line that these developments are incipient and isolated. But it also talks about the potentials that are foreseen and the problems needed to be resolved to arrive at the potentials. For example, the article discusses what AI can do in support of battle management as the field currently stands, what AI cannot do in support of battle management, and what AI cannot do *now* in support of battle management. The article also alludes to the DARPA Strategic Computing Program, which has been going on for several years now, and which has a major influence on artificial intelligence developments throughout DoD to the extent that the Program nearly always has representation at major AI forums.

The DAITC Hypermedia Laboratory, in taking its cue from the topic of information technologies in Defense, devised the following concerning the incorporation of *Hypermedia* in a Defense activity situation.

1. INTRODUCTION TO HYPERMEDIA

We have moved from the Industrial Age to the Information Age and with this new era comes the need to find means to handle all the information available to us. Within the last century technological advances made communicating with others in remote locations a real-time endeavor. Therefore, human knowledge was transferred at rates previously impossible. Then came the advent of the modern computer which has had a huge affect in bringing us forward into the Information Age. Now, we have a seemingly endless amount of information stored on computer in the form of various types of storage devices.

Hypermedia is a strong technique with which to handle all this information on diverse systems. The future of Information could be that of Hypermedia systems, a convergence of multiple media into one system. When I say *one system*, I am referring to one logical system as opposed to one physical system. The future of information will be workstations tied together by complex networks which will be capable of accessing remote systems as well as local machines. So to show you a glimpse at what could be, I'd like to take you on a trip into the not-so-far-off future.

2. SCENARIO

2.1 Part I, Crisis

The year is 1993 and country YY is heating up again. Within the last two weeks there has been a rash of violent acts in and around YYerun, the capital of YY. Recidivist terrorists have bombed a few restaurants and hotels, and have indiscriminately set off several car bombs. Prior to this rash of bombings, YY had become a relatively peaceful country. The conflict between YY and its neighbor, Qaroq, had ended over 2 years ago and a new peaceful leader in YY had emerged. The new leader and his staff had managed to turn the country around politically and socially. The major restructuring of the country is being hailed as an international model of cooperative action.

Because of YY's new found international role, they are hosting a week long conference and workshop for top United States military and governmental officials on large scale cooperative action sponsored by the

United States Government. The U.S. entourage is aware of the recent terrorist activities but the decision was made to go on with their plans of attending the conference and leave the internal policing of the streets to the local YYian guard. So off the attendees went from Dulles en route to Brussels for a refueling stop, than on to the conference in YYerun, YY.

Everything was progressing smoothly on the flight until the second leg when the terror of yesterday revisited the skyways -- *Yes, there are terrorists onboard waving plastic machine pistols with plastic bullets and C4 plastic explosive hand grenades with (you guessed it) plastic detonators.*

There was some panic from the passengers but the overwhelming feeling was one of hope, hope that their associates at home in the Pentagon could make use of their analysis system -- *hypermedia-based* -- to help end this catastrophe. The hijackers had taken over the radio and were already making demands for their "high priced hostages". They also demanded to be refueled in Athens, Greece. There was some resistance from the Athens Airport but the plane finally landed there without any difficulties.

So you are probably wondering what does this hypothetical event have to do with hypermedia?

2.2 Part II, The Analysis System -- Hypermedia-Based

Well, you happen to be on the team of analysts at the Pentagon assigned to recommend courses of action in the event of a national crisis of this nature. Your mission is to look at all the pertinent facts and information available to you and deduce what actions should be taken based on that knowledge.

The graveness of the situation calls for timely and informed analysis in hopes of taking action while the hijackers are in a stationary position. You immediately turn to your hypermedia workstation, your tool in your analysis of the situation.

So you're sitting at your workstation to initiate your crisis analysis plan of action. You open up several view windows on your very high resolution screen. each window color coded for keeping track of the views.

You begin by asking for all the information available that's coming off the wires on this particular hijacking. You enter your query by voice through a headset microphone, in natural language.

The system responds with a list of articles and messages from it's search. You pick the first one. Its text appears in one of the windows, and it has several words highlighted.

You choose the word "airport" and a picture of the Athens airport is displayed. This information has already been programmed in a high density storage medium as part of the crisis management program; the medium is transparent to the user, but may be a compact disc (CD) or video disc; the Athens airport information is displayed in color, contains both still and moving imagery, and has voice description in support of the visual information.

You go back to the document and choose "hijackers" which plays a recording of the hijackers voices from radio transmissions. This data was picked up and "written" into a disc that can later be erased for future situations. Simultaneously, in another window a picture is displayed that was taken of a hijacker in the cockpit of the plane. You send this information to the voice/video group in the Pentagon for analysis.

You return to the list of articles available to you from your initial query and choose another document. This document is about the layout of the airport and gives you the option of displaying the 3-D blueprints of the airport. These views can be turned to examine a chosen section of the airport from any direction. Following this you ask the system to bring up the information on the particular plane involved. A photo along with the dimensions and layout are displayed on the screen. Again, the image can be turned to examine the craft and its sections from any direction.

While you are viewing the plane layout a message appears on your screen indicating you have an incoming message from the voice/video analysis group. The message is voice, but the message text also appears in a screen window for later reference. You bring up the message on your screen for review. It consists of information and identification of the hijacker in the photo.

With the hijacker's name and organization you return to the query prompt in a selected window and ask for all information on the hijacker and the organization he belongs to. From the options presented to you after the search is completed you choose to view a video clipping of another hijacking incident he had been involved in back in the late 80's to get a feel for what he is willing and likely to do in this situation.

After traversing though more information you are confident you have an idea you want to run past a high ranking official in the Pentagon and one in the U.S. Embassy in YY. So you establish a connection with both, by voice, video, and computer system as available, over the transmission links as available (such as telephone, hard-wire, satellite) and begin to inform them of your analysis and recommendations, and show them supporting material through your videoconferencing link.

3. A REVIEW OF THE MEDIA IN THE HYPERMEDIA SYSTEM

The sources of the information are the ways the information is stored, whether permanently or temporarily. Sources are also the means that information is *transmitted*. The encompassing term used for those means is *media*. The linking of those media in a manner that lets the user arrange views of the information on a screen in a simultaneous, associative manner, and making use of text, imagery, sound, and color, is called *hypermedia*.

The following is a summation of available and emerging information technologies that could be incorporated into a hypermedia system. These technologies would be components of the hypermedia-based analysis system above.

3.1 OPTICAL DISCS (2)

Optical discs come in three standard sizes: 5.25 inch, 12 inch, and 14 inch. A 3.5 inch disc is beginning to appear on the market as an audio disc (music), and is referred to as a CD3. The disc has a reflective aluminum layer covered with a plastic coating. Data is "burned" into the disc in the form of minute pits and spaces. The data is picked up when a gallium arsenide laser beam is reflected by the aluminum layer and an optical pick-up senses the pits and the spaces. Benefits of optical disc technology are:

- Greater data density (up to 800MB on a 5.25 inch).
- Greater resistance to magnetic flux and temperature flux.
- Data life of 200 years.
- Stores text, pictures, graphics and sound.

But, searching an optical disc is 10 times slower than searching a magnetic disc.

3.1.1 CD-ROM

CD-ROM is Compact Disc, Read Only Memory. The CD-ROM must be industrially mastered, and the information on it cannot be changed. If there are changes in the information to be made, the entire disc must be re-mastered. CD-ROM has become standardized through drive manufacturers and the CD-ROM systems. The data format is standardized through HSFF (High Sierra File Format) from the International Standards Organization (ISO). As indicated by its designation, it can only be read by the user, not written on.

3.1.2 WORM

WORM stands for Write Once, Read Many. WORM allows users to write data once. Unfortunately,

WORM is not yet standardized, and a standard is not expected for two or three years. Another disadvantage is that it costs the user over \$3,000 for a WORM drive, as compared to about \$1,000 for a CD-ROM drive. A CD-WO is being developed jointly by Philips and Sony, and could open doors to write once CD's. Their standard is compatible with existing CD-ROM drives, and is around a year away.

3.1.3 CD-Audio

This medium is CD-ROM for recording and playing sound. Check with your teenager.

3.1.4 Re-Writable Optical Discs

This is a new technology. It is hoped to be on the market within 1-2 years. ISO is in process of coming up with an erasable media standard, but not yet for a data formatting standard. It is thought that erasable discs will compete with magnetic media, unlike other optical disc media. The reason for this is that it is expected that price per megabyte will be comparable to or less than magnetic media. A 5.25 inch disc will hold about 600 MB, and a 3.5 inch disc designed for laptop computers will hold 75 to a 100 MB.

There are three main optical disk technologies under development:

- Thermo magnetic optic, which is the most advanced.
- Phase change, which may be the cheapest to produce.
- Dye polymer.

ANSI recently accepted a composite file format standard.

3.2 Video Discs

Video disc technology incorporates a range of data presentations to the viewer that can include text, voice, sound and music, still and motion imagery, multiple tracks, e.g., bilingual language tracks, and various degrees of user interaction. The components of a video disc system are a drive and a television monitor. The three basic technologies are analog video discs, CD-I, and DVI, briefly described as follows:

3.2.1 Analog Video Discs

Analog video disc is a current technology. It is easily available on the market, although not to the extent that the CD-audio discs are. Analog video discs are most widely available as copies of motion picture films.

The analog video disc operates as follows: a video signal is picked up by a laser beam and is transmitted through a video overlay card, which mixes the computer graphics format of the video disc with the television processes, standardized by the National Television Standards Code (NSTC). The signal is then sent on to a television monitor capable of handling the combined signals.

3.2.2 CD-I

CD-I refers to Compact Disc-Interactive. This development is a product of Sony and Philips, coming out in the Fall of 1988 as a home product. It consists of a CD-I player attached to the home television set. According to an article appearing in Newsweek, October 3, 1988(3), the projected cost of a player is under \$1,000, and a number of publishers, including Time-Life, Grolier, and the Smithsonian Institution "are preparing hypermedia discs for the new machines."

The elements of the CD-I are:

- Provides for full motion video 1/3 screen.

- Very nice resolution and color capabilities.
- 750 colors for still video display.
- 256 colors for full motion video.
- 16 separate audio channels.
- Up to 16 hours of audio playing time.
- Stores up to 300,000 typed pages of text and data.
- Stores up to 7,000 still video pictures.
- Controlling hardware and software is proprietary.

3.2.3 DVI

Digital Video Interface is a new technology, being designed in part by the General Electric Laboratory. DVI technology uses digital compression and decompression technology as well as hardware and software products to bring interactive, full-motion video and audio to personal computers and consumer electronics. It functions as follows:

- Uses data compression/decompression at ratio of 120:1.
- Full motion over entire screen.
- Resolution equivalent to VHS recordings.
- 256 colors for full motion video.
- Uses standard CD-ROM player for program storage.
- Designed for 80286 under DOS.

DVI is expected to be available about the middle of 1990, at a system cost of about \$12,500.

3.3 Voice

Voice already has a voluminous technical literature. The DAITC has experimented with a voice driven program and found it to be feasible. In this case voice drive was established for the Common Command Language (CCL) development activity on an experimental basis, making use of a *non-continuous voice recognition* system. Our exploration of such an application was in essence another hypermedia vector for us. In a sense, it was a first step in the direction of having human-computer communication take place on a cognitive, human level. Its feasibility was shown in that voice could be used for standard command input to multiple information systems, and get results on the screen from those systems as output.

This testing also showed us the needs for making use of emerging voice technology. In the *non-continuous voice recognition* system environment, the system:

- o Must create voice profiles for every user.
- o Trains tokens for words in a phrase book.
- o Treats words as isolated utterances, with explicit pauses.

Voice technology is now emerging with *continuous voice recognition* systems. These systems recognize normal speaking, do not need to be "trained," and do not need a speaker profile in order to function. An import of this technology is the elimination of the need for word libraries, which otherwise take up large amounts of storage.

Digitized voice output is also possible.

3.4 Telecommunications

A major move in information science is to access remote information bases, as opposed to storing everything locally. The question is why duplicate information already stored, and why re-duplicate this information interminably.

The DAITC has a very strong Networking Laboratory dealing with interoperability and interconnectivity. The laboratory concerns itself with both internal and external networking, including globally, and deals in the communicating of disparate systems and their software.

Through networking telecommunications, the user can explore information sources, and collect needed information from them. The user can then arrange that information as needed. Once the collected and reformatted information has served its purpose, it can be stored as a specialized source of information, or disposed; the information can also be re-disseminated over the network systems.

3.4 Video Conferencing

Video conferencing permits a group of geographically distributed people, with a special interest, to convoke and confer electronically in a real-time interactive manner. This medium allows those people to apply greater focus to their interest without depending on having to get together in one place at a time that is convenient and possible for all.

This medium itself has a number of component media as a part of the conferencing process:

- o Text Mail - standard electronic mail (EM).
- o Voice Mail - digitized voice EM input and output.
- o Graphics Mail - the inclusion of graphics and imagery in EM.
- o Computer Conferencing - establishing links among the distributed participants via the computer and terminals, to do and accomplish whatever the computer is set up to do.
- o Audio Conferencing - Sound communications set up to carry the voices, music and noise of the distributed participants.
- o Augmented Audio Conferencing - to include for the distributed participants access to:
 - Slides and Printed Materials
 - Facsimile
 - Electronic Blackboard
 - Freeze-Frame Video
- o Videoconferencing - realtime, interactive, all-around televised conferencing of the distributed participants, to include any degree of the components above.

4. FUTURE DIRECTIONS

DAITC has all these capabilities and is working towards the vision of a hypermedia system to access all the information possible to further the strength of human knowledge.

A very small sample of some currently available Hypermedia tools are:

- NoteCards by Xerox
- HyperCard by Apple
- NeXT Machine by NeXT, Inc.

Artificial Intelligence is likely to add strength and intelligence to the Hypermedia field by using some of the following techniques when building Hypermedia applications:

- Learning
- Neural Networks
- Natural Language Interfacing
- Advanced Information Retrieval
- Intelligent front and back ends to knowledge systems

The article on AI in battle management includes the statement: "We Can't Manage Uncertainty." This paragraph continues: " There are several representations of uncertainty in the AI community, and there are spirited discussions each year about the relative merits of each, but it remains to be seen which of what combination of techniques is correct for battle management. . . ."

Although representation of uncertainty refers to an AI technique, the point is there is uncertainty about uncertainty. This is continues to be true for hypermedia systems, in any area of Defense criticality. But rapid-response tools for any kind of management could be provided, maintained, and enhanced; furthermore, each user/participant who would use them could add his or her knowledge to them for even greater usage by others.

REFERENCES

(1) "What AI Can Do for Battle Management: A Report of the First AAAI Workshop on AI Applications to Battle Management."

R. Peter Bonasso, *AI Magazine*: Vol. 9 No. 3, Fall 1988, pp. 77-83.

(2) "Eyeing the Optical Disk Market."

John Rizzo, -----: April 19, 1988, pp.20, 22.

(3) "Here Comes Hypermedia: Stereo Sound! Moving Pictures! Endless Data! It's Computer Heaven!"

Michael Rogers, *Newsweek*: October 3, 1988, pp. 44-45.

Ms. Edleson's interest in incorporating hypermedia systems into a crisis management environment stems from her background and training in such an environment: Ms. Edleson is a 1984 graduate of the U.S. Military Academy, a graduate of the Army Jungle Warfare School, and passed her Army Service as an officer in Explosives Ordnance Disposal.

This paper is based on a presentation made by Ms. Edleson at the Defense and Government Computer Graphics Conference, Washington, DC, October 1988. The paper has been augmented by Allan Kuhn, Director, DAITC Hypermedia Laboratory.

**HYPERTEXT
IN A RESOURCE SHARING
GREATER MACHINE LEVEL ENVIRONMENT**

Hypertext on the VAX

*Randy L. Bixby
Hypermedia Laboratory*

The staff of the Hypermedia Laboratory looked at PC-based hypertext functions and processes, and concluded that these functions could be imported into a resource sharing environment on the DoD Gate Information System (DGIS) machines, in this case, a VAX 11/780.

Hypertext on the VAX is an ASCII-based hypertext application in the super mini-computer environment. We developed it on an experimental basis as a possibility to provide hypertext capabilities to DGIS users who don't have bit-mapped terminals, but who could benefit from the contextual organization of information which is elemental to hypertext.

The first hypertext application prototyped on the VAX 11/780 is a small database containing an introduction to DROLS and its databases. Text for the database was manually entered using the vi editor on DGIS, although any other editor could be used. The text is broken into "chunks" of information to facilitate contextual, rather than linear, relationships between pieces of information. Each chunk is stored in a separately named file. Words or phrases within the chunks can be related contextually to other chunks (files) of information giving more detail about the word or phrase.

Hypertext on the VAX mimics bitmapped terminals by dividing the user's display screen into three windows. The lower window displays text files. The upper window displays the sequential path the user has followed through the database, listing files the user has selected (called BOOKMARKS). The center window continuously changes, appearing and disappearing as the user selects files or displays the index.

Hyper-Window System 10/03/88 09:21

```
-----
| BOOK MARKS
| DROLS Database (drolsl.hyp)
| Report Documentation Pages (reportdoc.hyp)
| Report Documentation Pages (drolsl.hyp)
| Report Documentation Pages (drols2.hyp)
|
| +-----+
|          Current_Technical_Reports_Database
|
| _____
|          CURRENT TECHNICAL REPORTS DATABASE (1 of 2) (ct1.hyp)
| INTRO | CURRENT TECHNICAL REPORTS DATABASE (2 of 2) (ct2.hyp)
|
| This +-----+
|
| Technical Reports Database (TR)
| Current Technical Reports Database (CF)
| Work Unit Information Database (WU)
|
| You may wish to consult various Reference Documents as you follow this course.
|
|-----
```

Contextual relationships are denoted by highlighting a term in reverse video on the screen display. Reverse video is programmed by enclosing the term or phrase between backslashes () in the text file. Every time a file is displayed, all characters which are enclosed in backslashes will appear in reverse video.

Links between highlighted terms and their related files are made through a linktable written in Prolog. The linking process is triggered when a database user selects a highlighted term by placing the cursor anywhere on the highlighted portion and hitting the RETURN key. The Prolog program performs several steps:

1. Is the term really highlighted? (The RETURN key could be hit when the cursor is not on a highlighted term). "Highlighted" is defined to be characters enclosed by backslashes. Prolog checks from the cursor position to the left and right, looking for backslashes. If the conditions are not met, an error message appears in the upper right corner of the top window, stating the link could not be made.
2. If the highlighting condition is met, Prolog compares the character string between the backslashes to items in the linktable. This is called pattern matching. If the character string "matches" an item in the linktable, that item appears in the middle window. If the character string matches more than one item, all matches will appear in the middle window. If no match is found, an error message appears in the top right corner of the upper window. (This would occur if a term is enclosed in backslashes in a text file, but no related file has been created, or no linking item has been entered in the linktable).
3. If a term appears in the middle window, RETURN will cause Prolog to retrieve the related file and display it in the lower window. If more than one term appears in the middle window, the user can cursor up and down in the list, and make his selection.

The user may jump around in the database. If he wishes to return to a previous screen display, he can jump from the lower (text) window to the upper (BOOKMARK) window, where highlighted terms he has selected are listed. He can move up and down in the list using the cursor keys, and select any of the terms in the list. The newly selected term will appear at the bottom of the BOOKMARK list, in addition to its original position in the list.

The user may also skip through the database by viewing the database index. This is done by entering the letter "t" (for "term"). A prompt appears at the bottom of the text window requesting entry of the desired term, and RETURN. The index appears in the middle window, showing terms closest alphabetically to the term the user has selected.

The user may request assistance by typing "?" at any time. The help window overlays the current screen display, showing a chart of cursor and screen movement keys.

Two other test hypertext applications have been developed on the VAX, using downloaded text. One consists of messages extracted from the RISKS Digest of USENET, and the other is the four evangelical books of the Bible. Additional applications are being considered, as well as alternate methods of inputting text, such as Optical Character Recognition (OCR).

The mainframe orientation of Hypertext on the VAX is ideally suited to networks and gateway systems such as DGIS, which has users around the world. Most hypertext applications are microcomputer or workstation-oriented, which means that only one person can use the application at a time, and the user must be in the same location as the machine on which the application resides. By placing an application on a multi-user computer such as a mainframe or minicomputer, many people can access the application at the same time, and the user and host computer can be miles apart.

Updates and enhancements to applications are done centrally by a database administrator. All users benefit from the changes automatically, without having to copy tapes or disks into their own machines. This

eliminates the need to create different versions for different types of microcomputers and workstations, and saves a lot of the time and effort involved in distributing software to a wide user group.

Database size is not a problem on a minicomputer as it may be on a microcomputer. Very large databases may be created, which in turn may be linked to other large databases. These databases need not even be located on the same machine, but can be on any machine accessible by a network.

DTIC plans to develop a hypertext template which will allow non-programmers to create their own applications. Users will be shown how to create or import text, divide the text into "chunks" or files, and how to create the links without having to know Prolog. Database creators can then set access privileges to their applications, so others on the network may use the database.

Hypertext on the VAX

--could provide *hypertext* capability to DGIS users. Scattered but related documents could be viewed as a single entity.

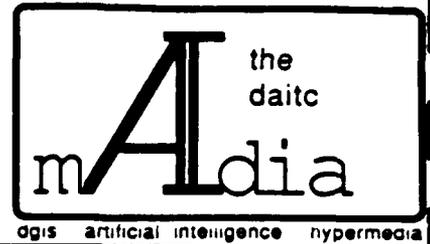
--could interface to DGIS resources to draw information to support the semantics of the documents: DGIS databases, DGIS CCL, DGIS access to external databases, access to information on DGIS networked computers.

--could apply AI techniques to provide flexibility: rules, algorithms to form and find aggregations, learning mechanisms to capture and personalize expertise.

Hypertext on the VAX has potential in many different areas--training, archives, budgeting, and simple information transfer. Hypertext is a technology whose time has come, and it will continue to grow in usefulness as more applications are developed.

Hyper-Window System 10/03/88 09:20

```
+-----+
| BOOK MARKS                                     |
| DROLS Database (drolsl.hyp)                   |
| Report Documentation Pages (reportdoc.hyp)    |
|                                               |
|               +-----+                       |
|               | KEY_ASSIGNMENTS              | |
|               |                             |
| _____    | ?                             | Help          |
| DD_FORM_1473, | n,p                         | Next, Previous Document |
| _REPO        | ESC                         | Cancel                |
| A 1473 is sub| w,b                         | Next, Previous word   |
| mitted       | +,-                         | Forward, Backward Link |
| describes wh| m                           | Select from bookmarks  |
| the          |                             | 73                    |
| done. It provides | Left,right cursor or h,l to move left,right | work wa
| as other numbers u| Up, down cursor or j,k to move left,right   | as well
| Technical Reports |                             | ut to th
|               +-----+                       |
+-----+
```



Hypermedia Lab. Program Note 5, October 1988

HYPERMEDIA & HYPERTEXT - Linking Up

Definition - Hypermedia is the linking of different information formats into a multi-dimensional network. The information formats may include text, graphics, videodisk, sound, or any other format. Another term is hypertext, which is the linking of bodies of text. What makes hypertext and hypermedia different from traditional computer information is that they are non-linear; that is, the reader doesn't have to follow straight through from the beginning to the end--he can make side trips for more information if he needs it, and completely bypass sections he is either familiar with, or that are unrelated to his current needs. Hypertext makes use of the computer's branching capability to move from one link to another. Automated hypertext systems are relatively new, but there are some manual equivalents which have been around a long time.

Manual Hypertext Systems - One example is the encyclopedia. As you read an article in an encyclopedia, you find references to related items which will have more information. These references are usually "non-linear"; that is, the other items or articles do not generally directly follow the one you are looking at. You may have to open another volume in the set. You may also be referred to a diagram or picture elsewhere in the encyclopedia. Another type of manual hypertext is the use of 3X5 cards for note taking. Note cards may be referenced to each other, and you might arrange them in a hierarchical fashion, such as grouping all the cards related to a single topic. An advantage of note cards is that their small size modularizes the notes into small chunks. The cards can be easily reorganized when new information suggests a restructuring of the notes, or some of the cards could be used in another paper at another time.

Hypertext Applications - The current definition of hypertext requires that navigation through the information is computer-supported. In the last few years, several working hypertext applications have been developed. Some are mainframe or minicomputer systems which allow multi-user access; others are microcomputer based for single users only. We are looking into both types of applications, and we have prototyped systems that run on the DGIS computer, and on PCs.

Concept of Hypermedia - The concept of hypermedia is quite simple: items on the screen are associated with objects in a database, and links are provided between these objects. The user can move from one object to another by following those links. Linked objects are identified in the text or database in one of several fashions, usually highlighting terms in the text or establishing graphic representations of terms or concepts. These representations are called icons or buttons.

Sample Hypertext System - The sample hypertext system uses the highlighting technique. Selecting any of the terms you see highlighted activates the link to further information, which is then displayed on the screen. The user can navigate through the database by selecting highlighted terms in the text, or by selecting standard options displayed at the bottom of each screen, or by going to the index, where all the subject terms in the database are listed.

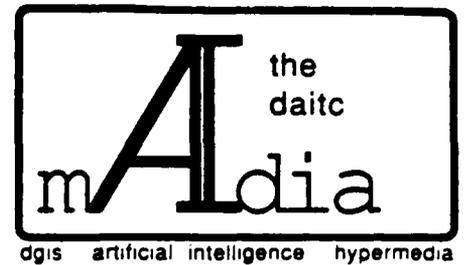
Applications for Defense Technical Information Center (DTIC) - At DTIC we first became interested in hypermedia in relation to the Common Command Language project. Our first application was the development of a "desktop" environment on the Sun workstation. The Sun is a bitmapped system, which allows graphics displays and multiple windows. The desktop uses menu bars and icons, which are graphic representations of commands or functions. If you have ever used a Macintosh microcomputer, you're familiar with icons and menu bars. The Sun also uses a mouse, so all you have to do is point and click to open and close windows or select options. The desktop on the Sun is being developed to determine the feasibility of this type of system. If it proves practical, it will be modified and ported into DGIS so it is available to users who have bitmapped systems. We are also developing a hypertext system on the DGIS computer for non-bitmapped systems. This system will offer windowing and highlighting for textual databases in the character-based environment.

The Future - Hypermedia and Hypertext offer exciting possibilities for the linking and manipulation of information resources in different formats. We will continue to examine applications which will make the information-seeking process simpler and more adaptable for our users. We are especially interested in developing applications linking text with CD-ROM, voice-activation and audio resources, graphics, and videodisk.

The prime goal of the Hypermedia Lab. is to create AI-based programs that make the human-machine interface more human in the machine's responses.

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER (DAITC) ALEXANDRIA, VIRGINIA

Hypermedia Lab. Program Note 6, October 1988
Voice Lab. joint note



VOICE & the gateway computer system

Speak !

A major underlying dream of most computer developers is to create a computer system capable of communication with humans on an intelligible level. This dream encapsulates several research areas under the umbrella of Artificial Intelligence and Hypermedia. The most efficient way humans communicate is through mutual oral or written information exchanges. The information passed between humans is usually semantically meaningful. Thus, for a computer to successfully "communicate" in the same realm as a human, it must be capable of natural language understanding, speech recognition and speech synthesis, to name a few relevant areas of study in Artificial Intelligence.

The Hypermedia and Voice Labs of the DAITC recognized this research area of computer-human communication, and have developed an experimental prototype Voice-driven CCL system. This joint effort combines the existing DAITC phonological technology and the CCL system to produce an Artificial Intelligence driven Hypermedia environment to access heterogeneous database systems. The development shows that speech and textual input could feasibly access remote databases online, on discs, and on other storage media, through various networks to produce a Hypermedia system. And, the user input and system output is controlled by the Artificial Intelligence engine of the Common Command Language System.

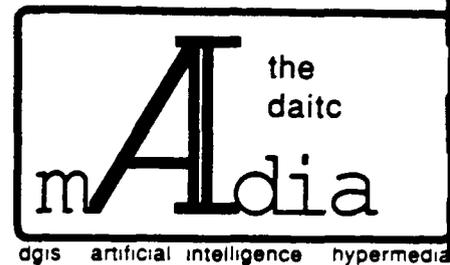
The actual path of communication in the Voice-driven CCL system begins with the user input. The current DAITC phonological technology consists of a Kurzweil AI Voice box that accepts trained speech input from the user. The Kurzweil box then translates the input and sends the ASCII equivalent output to the PC for processing. The PC in turn accepts this information as if it were entered from the keyboard as input. So to run CCL using voice input the user enters the voice commands to logon to the DGIS system and run CCL. Once in CCL the user enters, by voice, the preferred database system to search on. Then within the database system the user is free to use any of the CCL commands to search and display terms in the database. The CCL system then responds to the voice input with normal CCL screen output.

The available voice input search terms are limited by the user trained vocabulary library controlled by the Kurzweil software and hardware currently resident at the DAITC. But new speech recognition systems are now available that accept input from unlimited users without individual user voice training.

The DAITC Voice-driven CCL prototype demonstrates the feasibility of using other mediums for input in the DGIS environment. This obviously leads one to wonder what other technological enhancements are capable for integration into the DGIS environment or eventually the STINET workstation environment. All in all, another step has been taken at the DAITC towards the dream of an intelligent communicating computer.

The prime goal of the Hypermedia Lab. is to create AI-based programs that make the human-machine interface more human in the machine's responses.

**DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER (DAITC)
ALEXANDRIA, VIRGINIA**



minicomputer * HYPERTEXT * multiuser

Hypermedia for ASCII

Hypertext on the VAX is an ASCII-based hypertext application in the mini-computer environment. It was developed to provide hypertext capabilities to DGIS users who don't have bit-mapped terminals, but who could benefit from the contextual organization of information which is elemental to hypertext.

The first hypertext application prototyped on the VAX 11/780 is a small database containing an introduction to DROLS and its databases. Text for the database was manually entered using the vi editor on DGIS. The text is broken into chunks of information stored in separately named files. Words or phrases within the chunks can be related contextually to other chunks (files) of information giving more detail about the word or phrase. The relationship of a term to further information is denoted by highlighting the term in reverse video on the screen display. Users of the database can select a highlighted term using cursor keys on the keyboard to move up and down, left and right, on the screen, and then hitting the RETURN key. This follows the link from the term to its related document. Links are programmed using the pattern-matching properties of PROLOG computer language, resident on the VAX.

Hypertext on the VAX mimics bitmapped terminals by dividing the user's display screen into three windows. The lower window displays text files. The upper window displays the sequential path the user has followed through the database, listing highlighted terms the user has selected (called BOOKMARKS). Accumulated terms scroll the list from bottom to top of the window. The center window echoes the most recently-selected highlighted term, the term index, or displays optional paths the user may follow.

The user may jump from one window to another. Cursor keys may be used to move around in all three windows. The user may also skip through the database by entering the letter "t" at any time. A prompt appears requesting entry of the desired term, and RETURN. The index appears in the middle window, showing terms closest alphabetically to the term the user has selected. Help is available at any time by typing "?". The help window overlays the current screen display, and gives a chart showing how to move around in the database.

Two other test hypertext applications have been developed on the VAX, using downloaded text. One consists of messages extracted from the Risk Digest of USENET, and the other is a portion of the Bible. Additional applications are being considered, as well as alternate methods of inputting text, such as Optical Character Recognition (OCR).

The main-frame orientation of Hypertext on the VAX is ideally suited to networks and gateway systems such as DGIS, which has users around the world. Most hypertext applications are microcomputer or workstation-oriented, which means that only one person can use the application at a time, and the user must be in the same location as the machine on which the application resides. By placing an application on a multi-user computer such as a mainframe or minicomputer, many people can access the application at the same time, and the user and host computer can be miles apart.

Updates and enhancements are done centrally, and all users benefit from the changes automatically, without having to copy tapes or disks into their own machines. This eliminates the need to create different versions for different types of microcomputers and workstations, and saves a lot of the time and effort involved in distributing software to a wide user group.

Database size is not a problem on a minicomputer as it may be on a microcomputer. Very large databases may be created, which in turn may be linked to other large databases. These databases need not even be located on the same machine, but can be on any machine accessible by a network.

Hypertext is an expanding technology for the times, and will continue to grow in usefulness as more applications are developed.

See also Hypermedia Lab. Program Note 5, Oct 88, HYPERMEDIA & HYPERTEXT.

The prime goal of the Hypermedia Lab. is to create AI-based programs that make the human-machine interface more human in the machine's responses.

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER (DAITC)
ALEXANDRIA, VIRGINIA

S U M M A R Y

SUMMARY

&

A PRESENTATIONAL OVERVIEW OF THE DAITC HYPERMEDIA LABORATORY ACTIVITIES

First presented at the Defense & Government Computer Graphics Conference

(DGC88-East) -- October 1988 -- Washington, D.C.

Allan Kuhn

Director, DAITC Hypermedia Laboratory

December 1988

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER

DAITC

SPONSORS - OSD, JOINT STAFF, DLA

PURPOSE - CENTRALIZED FLEXIBLE ENVIRONMENT

- DOD IRM REQUIREMENTS
- TIMELY & ECONOMICAL MANNER

MAJOR FUNCTIONS

- TEST & EVALUATE NEW INFORMATION TECHNOLOGIES
- IDENTIFY OFF-THE-SHELF TECHNOLOGY
- DEVELOP TURNKEY PROTOTYPE/PILOT SYSTEMS
- DEVELOP INTEROPERABILITY SOLUTIONS
- PERFORM SYSTEMS INTEGRATION
- PROVIDE ASSISTANCE DURING AIS DEVELOPMENT LIFE CYCLE
- DEVELOP INTERACTIVE MEDIA TRAINING COURSEWARE

In 1988 the development of the DAITC organization and mission was sponsored jointly by the Office of the Secretary of Defense (OSD), the OSD Joint Staff, and the Defense Logistics Agency (DLA).

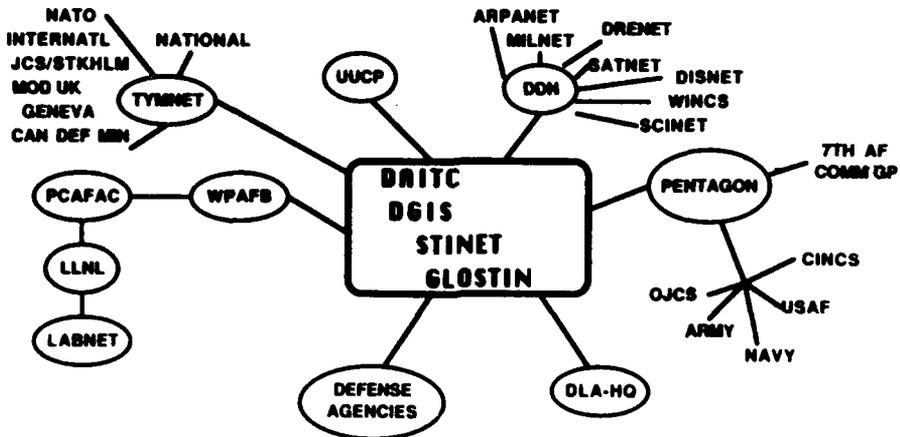
The mission of the DAITC was to act as a centralized information applications development environment to serve the information resource management (IRM) needs of the Department of Defense (DoD). Having this environment would help DoD organizations meet their information requirements in a timely and economical manner. The Center would be able to accelerate the introduction and use of new technologies for information handling, and because of that be the focal point for the development of new systems.

The Director of the DAITC is Ms. Gladys A. Cotter, one of its primary organizers. Ms. Cotter reports to the Administrator, Defense Technical Information Center, which in turn is a primary field activity of the Defense Logistics Agency.

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER

DAITC

NETWORKING * INTEROPERABILITY * INTEGRATION

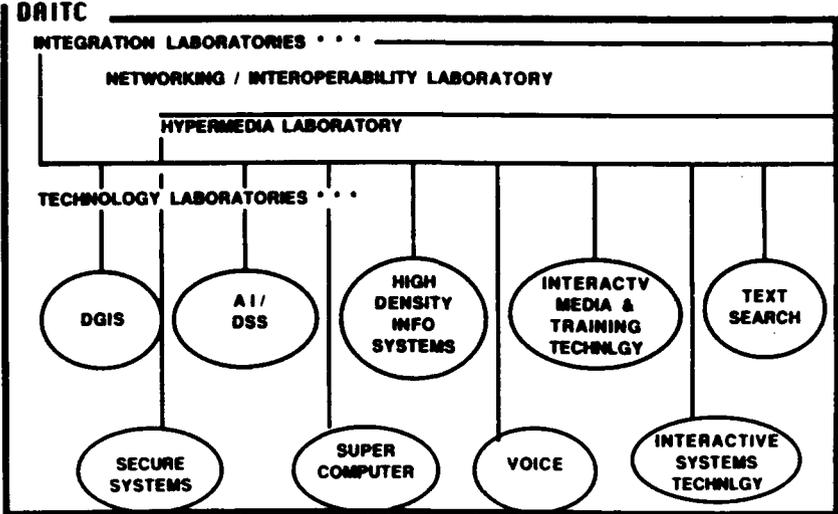


A critical key in making information available is establishing networking systems to get the information user to the information. The components for doing this are: the connection (networking) with systems and databases, creating the conditions for intersystem compatibility (or interoperability), and thereby achieving an environment in which the systems are integrated in the view of the user.

In the DAITC's world view, the core system is the Defense Gateway Information System (DGIS), a streamlined system which acts as the portal for taking the user out to the world. We expect DGIS to develop into the Scientific and Technical Information Network (STINET), which will give access not only to the major Government and commercial information systems, but also to the databases throughout DoD such as those in Information Analysis Centers (IACs), DoD Technical Libraries, and databases located in scientific, engineering, and management organizations and offices. Global STINET (GLOSTIN) envisions world-wide information access.

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER

DAITC



There are two categories of laboratories in the DAITC: Integration Labs and Technology Labs. There are two integration labs: the Networking and Interoperability Laboratory, and the Hypermedia Laboratory.

The integration labs are concerned with the potentials of integrating the developments that come out of the individual technology laboratories. The Networking and Interoperability Lab is the key to all technology integrations, including those dealt with by the Hypermedia Lab, with its concern about the compatibility of hardware and software systems and communications. The primary difference between the two integration labs is that the Networking and Interoperability Lab concerns itself with the accessing of information; the Hypermedia Lab is concerned about the presentation of information. The Hypermedia Lab views the hypermedia system as giving the user a *contextual* and *graphics* view of information, as compared to the traditional linear, textual view.

The individual technology labs shown are examples of areas of either development or interest at the DAITC

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER

HYPERMEDIA **DAITC** **LABORATORY**



ARTIFICIAL INTELLIGENCE

BITMAP SYSTEMS

WINDOW TECHNOLOGY
SIMULTANEOUS WINDOW DISPLAYS
SIMULTANEOUS TASKING / DISPLAYS
SIMULTANEOUS DISPARATE TASKING / DISPLAYS
COLOR / SOUND

NETWORKING

MEDIA

We find ourselves working with four elements that create a hypermedia system.

ARTIFICIAL INTELLIGENCE - was the initial activity of the Hypermedia Lab; we see AI becoming submerged in systems, to serve as the interface engine to the multiplicity of databases and their storage sources.

BITMAP SYSTEMS - is the vehicle for our stepping into the hypermedia realm. It is the capabilities of the bitmap, or graphics, system that allow putting information into contextual displays.

NETWORKING - is the gateway to geographically distributed information sources, whether within the organization or throughout the world.

MEDIA - is the storage means of information -- online, and optical discs: CD-ROM, WORM, video-interactive; media also include variable sources of information such as floppy disks, electronic mail, digitized voice, video conferenceing; extended media include *stimuli information* - color, audio, and imagery - both still and moving.

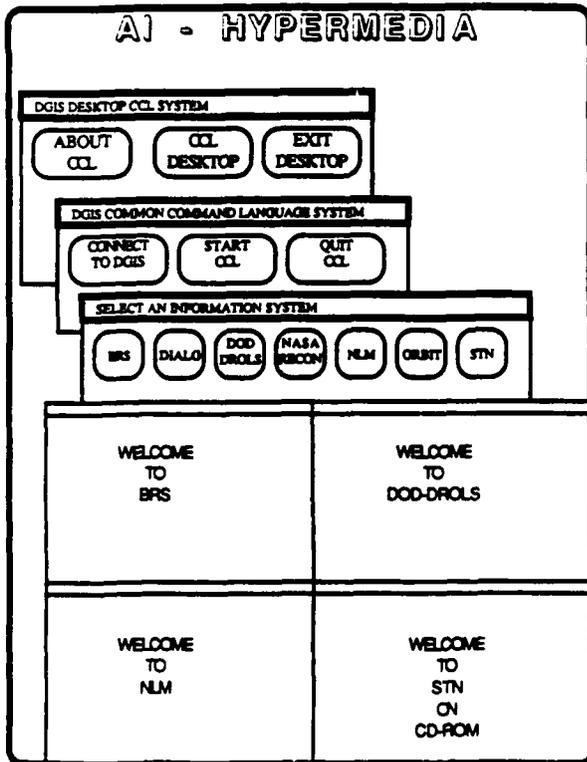
DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER

HYPERMEDIA DAITC LABORATORY

ACTIVITIES

- COMMON COMMAND LANGUAGE SYSTEM
- ARTIFICIAL INTELLIGENCE
- QUERY BUILDER / PREPROCESSOR ES
- AI/ES/OODBMS
- DESKTOP ENVIRONMENTS
- HYPERMEDIA/BITMAP SYSTEMS
- VOICE
- PHONOLOGY/AI
- NETWORK MGT ES
- HYPERMEDIA/BITMAP
- OBJECT ORIENTED ENVIRONMENTS
- AI/HYPERMEDIA
- HYPertext IN MAINFRAME ENVIRONMENT
- HYPERMEDIA LINKING / CHARACTER-BASED
- NATURAL LANGUAGE ENVIRONMENT
- ARTIFICIAL INTELLIGENCE
- GENETIC ALGORITHM ENVIRONMENT
- AI/HYPERMEDIA

X-WINDOWS
GRAPHICS PROGRAMS NETWORKING



This was our first applied concept of a hypermedia-based DGIS, using the Common Command Language development as the vehicle -- a *desktop system* making use of a bitmap system's capabilities of icons, multiwindowing, and multitasking. Color would be an enhancement to the displays.

We were able to look at quite a number of technologies to varying degrees during 1988.

The *Common Command Language System (CCLS)* was our first development activity. The CCLS is an artificial intelligence based system to access heterogeneous information systems in a standard manner through the DGIS. It is written in PROLOG, and makes use of knowledge base and black-board architectures.

The *Query Builder / Preprocessor Expert System (QBES)* will be concerned with query structuring using expert systems principles, and integrated into CCLS.

We began looking at *Desktop Environments* with SUN workstations. Our first rapid prototype concerned using icons and the multiwindowing capabilities for CCLS. Our second rapid prototype made use of an *Object Oriented* programming package for setting up a desktop hypermedia system for CCLS. The value of the package was shrinking desktop prototyping from weeks to days.

We experimented with *Voice* application, again using our CCLS development, to test its feasibility. At the time our equipment consisted of the *non-continuous voice recognition* technology. We were very pleased with this experiment because it showed the feasibility of using voice commands for invoking DGIS functions, searching information systems, and getting results. This also showed a closer integration of the human and machine, making the machine conform to human functions.

With the rapid expansion and networking of systems and equipment at the DAITC (a VAX, a PY-RAMID, a GOULD, communications equipment, LANs, SYMBOLICS, SUNS, *et al.*), the potential was seen for a hypermedia-based *Network Management Expert System*. This system would be AI-run, and handle text and images. Such a system would not only document the DAITC network, but also help monitor it.

It is an extremely easy and even essential step to incorporate *Object Oriented Programming Environments* into hypermedia, not only for their more rapid development capabilities, but also for their capacity to handle image objects in addition to text objects. A major benefit of OO programming that we have experienced is being able to shorten window-based programming from weeks to days.

One of our strictly experimental areas was developing a *hypertext function in a mainframe environment*. We put up a system on a VAX 11/780 machine, programmed in PROLOG, and operating in a character-based UNIX environment. There were threefold results: one was looking at hypertext concepts and processes, the second was creating a hypermedia process that functioned in ASCII, and the third was putting up a hypertext system as a shared resource.

When we first began working with artificial intelligence, a natural area of concern was *natural language*. We could see, for example, the DGIS Common Command Language eventually migrating to natural language. Another foreseeable area is natural language interfaces. We have not yet begun an intensive application of natural language processes, however, but more toward simple applications, such as making the use of English, rather than system commands, integral to DGIS operations.

The *genetic algorithm environment* concerns the refinement of information search results through the application of chromosomal principles, *i. e.*, only the fittest survive. Results are cycled through performance or fitness criteria, until a set of results conforming to those criteria is achieved. At this point, we have only been looking at the principles of genetic algorithms.

X-WINDOWS came to our attention toward the end of 1988. Because *X* is a software program for distributed bitmap system information handling, we immediately saw it as a possibility for a truly hypermedia network system, sharing the resources of DGIS/STINET in a contextual manner.

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER

HYPERMEDIA DAITC LABORATORY



RESPOND TO THE SURGE IN HYPERMEDIA TECHNOLOGY

IDENTIFY TECHNOLOGIES THAT INTEGRATE INTO THE HYPERMEDIA CONCEPT

MAINTAIN AN EMERGING TECHNOLOGIES WATCH

IDENTIFY THE HYPERMEDIA TECHNOLOGIES THAT RELATE TO DAITC ACTIVITIES

INTEGRATE THE COMPONENT DAITC TECHNOLOGIES

Originally I was going to title this "*objectives*" or "*goals*" or some such thing, but then I realized the flatness of these bureaucratic things when talking about a subject as wide open and enthusiastic as *HYPERMEDIA*. I then came up with *VISION* --

SURGE - If the DAITC is to act as an edge in information handling, then we have to acknowledge and make use of the developments that are exploding throughout the information world --therefore, be responsive and responsible to hypermedia.

IDENTIFY - With the technologies available at the DAITC -- optical disk, artificial intelligence, voice, interoperability and networking -- these things naturally lend themselves to full-fledged integration into hypermedia systems. Our first technologies identified for integration was Voice and the DGIS Common Command Language, on an experimental and feasibility basis.

WATCH - I feel the most important element is the *Technology Watch*. This requires having people who are dedicated and intensely interested in the technologies, and follow all the sources -- the literature, peer groups, the network announcements and exchanges -- for developing and emerging technologies.

RELATE - Relating hypermedia technologies to DAITC activities is highly open-ended, but what this really means is seeing what's new and arriving and seeing how they can be integrated into the DAITC information handling programs.

INTEGRATE - In both making use of technologies already at the DAITC and bringing in new technologies, the integration of those technologies becomes an activity of the DAITC as a leading edge organization.

This is an attempt to formulate a hypermedia-system strategy for the DAITC, under the activities of the Hypermedia Laboratory.

1. Because DGIS is the core system of the DAITC, DGIS is the baseline system for the development of adjunct and supporting hypermedia systems. These systems should be developed for those in the DoD information community that make use of bitmap systems as the tools of their information activity. As the costs of bitmap systems decrease, their use will expand throughout DoD. An example of lowering costs is that of a Macintosh II, ranging from \$12,000 to \$30,000 depending on what is included, as compared with the older SUN 3 workstations, ranging from \$60,000 to \$100,000 depending on hardware and software configurations. [Note: SUN Microsystems is now developing much lower-costing systems that will be just as powerful as the SUN 3s if not more.]

2. These are developments and implementations that integrate the development activities of the individual labs of the DAITC.

3. The final goal is to present the DAITC as an organization that is seriously involved with and making contributions to

HYPERMEDIA

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER

HYPERMEDIA DAITC LABORATORY

GOALS & STRATEGY

1. RESPONSIBLE FOR DAITC AI AND HYPERMEDIA PROGRAMS
IN SUPPORT OF DGIS / STINET
2. RESPONSIBLE FOR ACTING ON HYPERMEDIA DEVELOPMENTS
AND ACTIVITIES THAT INTERRELATE DAITC LABS
3. RESPONSIBLE FOR ESTABLISHING A DAITC PRESENCE IN
HYPERMEDIA IN THE INFORMATION COMMUNITY

BRANCHING * LINKING * INTEGRATING

HYPERMEDIA LABORATORY GOALS FOR 1989

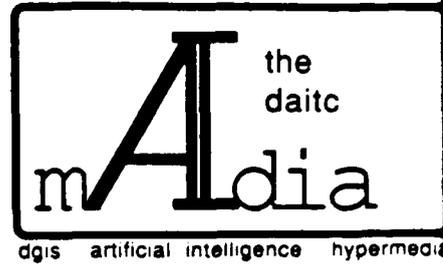
Project Goals

1. Completion of DTIC Major Project for DGIS Common Command Language System (CCLS) Design. This development will provide a standard, universal way of searching five major information systems: DIALOG, BRS, ORBIT, DROLS, NASA; a sixth information source, ENERGY, is now on DIALOG. The CCLS will then provide the frame for applying CCL to other systems accessible through DGIS.
2. Initiation of DTIC Major Project for DGIS Query Builder Expert System (QBES) Design. QBES will make use of CCLS, AI and ES concepts in query formulation for addressing multiple databases. This project initiates a potential Diverse Database Query Expert System which could involve distributed AI and ES (DAI) concepts.
3. Formalization and development of AMC Foreign Market Analysis System (FMAS) gateway interface design. The objective of the FMAS is to automate the identification and analysis of foreign commodity and technology market opportunities for the International Cooperative Programs Office (AMCICP) of the Army Material Command (AMC). The gateway interface will integrate C language and AI-based programs for pre-access, connecting, searching and retrieving, and support features providing search options, accounting, tutorial modes, *et al.*

Laboratory Goals

1. Incorporate object-oriented programming (OOP) and database management systems (OODBMS) into Hypermedia Laboratory developments.
2. Integrate and interface the advanced developments of CCLS, the Directory of Resources, and any future systems applicable to DGIS, such as the possibility of online thesauri (e.g., DROLS DRIT).
3. Continue review and implementation of X-WINDOWS into the DAITC environment, prior to hypermedia-based networking developments.
4. Incorporate C language based Expert System shells into the Hypermedia Laboratory and DAITC development environments; this includes evaluation and assessment prior to acquisition.
5. Expand on artificial intelligence and expert system -based developments and implementations; this will occur initially with the FMAS gateway developments, with implications for importing into the DGIS.
6. Establish a link with the Massachusetts Institute of Technology concerning the importation of AI and ES -based concepts and programs for handling information; this will be done primarily with its Laboratory for Information and Decision Systems, initiated with the FMAS project.
7. Study and initiate hypermedia programs that are technical library oriented.
8. Review the development of media machines and their new and emerging technological capabilities.
9. Continue the technology watch for relevant hypermedia integrations.
10. Continue the education of DAITC and DTIC staffs in hypermedia.
11. Expand DAITC technical staff in *across-the-walls* hypermedia activity, development, and implementation.

APPENDICES



PRESENTATIONS & PUBLICATIONS 1988

DAITC Hypermedia Laboratory Presentations and Publications - 1988

PRESENTATIONS

RIAO 88: User-Oriented Content-Based Text and Image Handling, Massachusetts Institute of Technology, Cambridge, MA, March 21-24 1988

Session: "Knowledge Based Systems."

National Online Meeting, New York, May 10-12, 1988

Session: "AI and Expert Systems Aiding Online."

Special Libraries Association (SLA) 79th Annual Conference, Denver, CO, June 11-16, 1988

Session: "Expert Systems in Libraries."

Defense & Government Computer Graphics Conference, Washington, DC, October 24-28, 1988

Session: "Hypermedia."

DTIC '88 Annual Users Conference, Alexandria, VA, October 31 - November 3, 1988

Session: "DTIC Showcase: Special Projects."

12th International Online Information Meeting, London, England, December 6-8, 1988

Session: "Intelligent Gateways: Systems."

PUBLICATIONS

Kuhn, Allan D., Randy L. Bixby, and Duc T. Tran. "DoD Gateway Information Systems (DGIS) Common Command Language: The Decision for Artificial Intelligence." Proceedings: RIAO 88 Conference - User-Oriented Content-Based Text and Image Handling Systems, Massachusetts Institute of Technology, Cambridge, MA, March 21-24 1988; CID, Paris, France, Mars 1988, pp. 863-882. Reprint available as AD-A199 215.

Kuhn, Allan D. "DoD Gateway Information System (DGIS) Common Command Language: A Summary of the First Prototyping and the Decision for Artificial Intelligence." Proceedings of the Ninth National Online Meeting, New York, May 10-12, 1988; Learned Information, Inc., Medford, NJ, 1988, pp. 169-183. Complete version available as AD-185 950, by Allan D. Kuhn, Randy L. Bixby, and Duc T. Tran.

Kuhn, Allan D., and Gladys A. Cotter. "The DoD Gateway Information System (DGIS): The Department of Defense Microcomputer User's Gateway to The World." Microcomputers for Information Management: An International Journal for Library and Information Services, Ablex Publishing Corp., Norwood, NJ, 5 (2) June 1988, pp. 73-92.

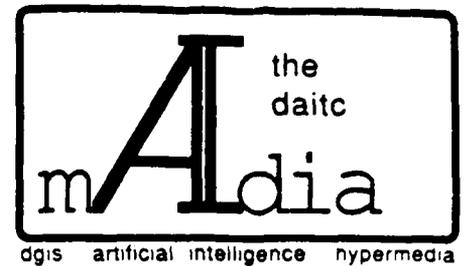
Kuhn, Allan D. "DoD Gateway Information System (DGIS): The Development Toward Artificial Intelligence and Hypermedia in Common Command Language." ONLINE INFORMATION 88: 12th International Online Information Meeting, Proceedings; 6-8 December 1988; Learned Information Ltd., Oxford, England, 1988, pp. 691-704.

The prime goal of the Hypermedia Lab. is to create AI-based programs that make the human-machine interface more human in the machine's responses.

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER (DAITC)

ALEXANDRIA, VIRGINIA

DAITC HYPERMEDIA LABORATORY
FORERUNNER PUBLICATIONS 1987



Predecessor Development Activity Documentation

The following documentation resulted from development activity that preceded the DAITC and the Hypermedia Laboratory. This development and its documentation laid the foundation, however, for the establishment of the Lab. As indicated by the citations, the developments are done under the sponsorship of the Defense Technical Information, Center (DTIC), one of the sponsors of the DAITC.

DTIC AD-A181 100

Toward and Artificial Intelligence Environment for DTIC: Staffing Qualification Criteria for AI Application Development. Allan D. Kuhn and Duc T. Tran. February 1987, DTIC AI Foundational Series No. 1.

DTIC AD-A181 101

Artificial Intelligence Developments Re: DoD Gateway Information System (DGIS) & Defense Applied Information Technology Center (DAITC). Allan D. Kuhn. February 1987, DTIC AI Foundational Series No. 2.

DTIC AD-A181 103

Toward an Artificial Intelligence Environment for DTIC: Proposed Tasks, Recommended Configurations, Projected Start-Up Costs. Allan D. Kuhn. May 1987, DTIC AI Foundational Series No. 3.

DTIC AD-A185 950

DoD Gateway Information System (DGIS) Common Command Language: The First Prototyping and the Decision for Artificial Intelligence. Allan D. Kuhn, Randy L. Bixby, and Duc T. Tran. August 1987, DTIC AI Foundational Series No. 4, DTIC CCL Report No. 1.

DTIC AD-A185 851

DoD Gateway Information System (DGIS): Common Command Language Mapping. Randy L. Bixby. October 1987, DTIC CCL Report No. 2.

DTIC AD-A186 150

DoD Gateway Information Systems (DGIS) Common Command Language: PROLOG Knowledge Base Profile. Duc T. Tran. October 1987, DTIC AI Foundational Series No. 5, DTIC CCL Report No. 3. Also published in *Quintus* ?- *consult(user)user.:The Quintus Newsletter*, as "Common Command Language: A Standard Language for Bibliographic Databases." Quintus Computer Systems, Inc., Mountain View, CA, September 1987.*

* Because of the *Quintus* article, a request was received from the University of Edinburg, Centre for Cognitive Science, for more information about the CCL development and the use of PROLOG. The University of Edinburgh is a major center of PROLOG development. Additionally, the National Information Standards Organization (NISO), National Bureau of Standards (NBS), requested the full documentation on the DTIC CCL development.

The prime goal of the Hypermedia Lab. is to create AI-based programs that make the human-machine interface more human in the machine's responses.

DEFENSE APPLIED INFORMATION TECHNOLOGY CENTER (DAITC)
ALEXANDRIA, VIRGINIA

Personal Acknowledgements
1988

The DAITC Hypermedia Laboratory was able to emerge only with the help of many, many people. This is to thank the following people, who have had an involvement in establishing the Hypermedia Laboratory, either directly or in supporting roles.

Ms. Gladys A. Cotter, Director, DAITC -- Ms. Cotter lay the foundation for the DoD Gateway Information System (DGIS), served as Chief, Information Systems and Standards Division of the Defense Technical Information Center (DTIC), and was a founder of the Defense Applied Technology Information Center (DAITC) prior to her becoming its Director. Ms. Cotter maintains a constant awareness of developments in information technologies, and she has extensive qualities in marshalling resources and using those resources to create an environment for making advances in information technology.

Dr. Esther Horne -- Dr. Horne was formerly at the School of Library and Information Science of The Catholic University of America, Washington, D.C., and is now at Chatham College, Pittsburgh, Penn. Dr. Horne had two fundamental approaches to her students, of whom I was one (1982-1984): she had an institutional approach which was to make her students aware of new and emerging information technologies, in addition to traditional and current; she had a personal approach which was to serve her students as extensively as possible, including when necessary "saving the patient." Dr. Horne helped lay the conceptual groundwork for the Hypermedia Laboratory through her classes and through an amount of consultation freely given. I hope this document shows that the patient was worth the saving.

Mr. Donald Bell -- Mr. Bell, of TRESP Associates, Alexandria, Va., served as the contractual DAITC Facilities Manager. Mr. Bell has extensive management capabilities with which he not only manages the personnel and plant but also manages in an exemplary and highly motivated manner the overall working environment. It is thanks to Mr. Bell that the physical establishment of the Hypermedia Laboratory was achieved.

Hypermedia Laboratory Staff

Mr. Duc T. Tran, Consultant to CDC -- Mr. Tran is a very highly knowledgeable computer scientist with extensive expertise in artificial intelligence. Mr. Duc strives for excellence in his work, and in doing so maintains a strong awareness of emerging and developing technologies for applications in hypermedia. He has made many suggestions and contributions to developments that have been extremely beneficial to the Hypermedia Laboratory activities, in addition to being responsible for technical programming and implementation. Mr. Duc is the Chief technical person of the Hypermedia Laboratory.

Ms. Randy L Bixby, DTIC -- Ms. Bixby came to DTIC with a Master's in Library Science, and with this background has rapidly developed an expertise for researching and establishing requirements for information handling. She is doing this for the DGIS Common Command Language System, and is initiating hypermedia-based information handling systems for the technical library environment.

Ms. Brenda A. Edleson, Control Data Corp. (CDC) -- Ms. Edleson joined the Hypermedia Laboratory in April 1988, bringing with her experience in expert systems and artificial intelligence programming, in addition to extensive and varied 3rd generation level program implementations. She also brought with her the unusual aspect of being a West Point woman graduate (1984), followed by a short term in the U. S. Army. After researching and completing a variety of AI and hypermedia tasks in the Laboratory, Ms. Edleson departed in November 1988 on an entrepreneurial small business consultant endeavor.

Additional REVIEW Contributors

Ms. Judy F. Hunter, DTIC -- Ms. Hunter is Director of the the DAITC Networking and Interoperability Laboratory. She has brought with her extensive expertise and skills in networking and systems compatibility. Her background includes having been responsible at one time for six computer systems, three on the east coast and three on the west coast. Since coming to DTIC and the DAITC in 1987, she has assured DAITC communications locally, nationwide, and worldwide.

Mr. Richard S. Marcus, Chief, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology -- Mr. Marcus is widely known and well published throughout the information world for his CONIT development work. Mr. Marcus has developed concepts and applications for information handling and modelling that are of great interest for the DGIS environment.

Mr. Albert Yanez, Centre de Documentation de l'Armement (CEDOCAR), Paris, France -- Mr. Yanez is Assistant to the Director of CEDOCAR. CEDOCAR is the French Ministry of Defense equivalent to the U. S. DTIC. Mr. Yanez's interests in the advancement of information handling parallel our own, and because of that Mr. Yanez and DTIC staff have exchanged visits and information, both bilaterally and at the NATO level.

DAITC and DGIS Technical Staff

The DAITC and DGIS technical support is provided under contract with Control Data Corp. (CDC).

Mr. Curtis Generous, Consultant to Control Data Corp. (CDC) -- Mr. Generous is a computer scientist *sans pareil* of the highest level. He has been with the DoD Gateway Information System (DGIS) development since 1985. He is currently the head technical person for the DAITC computers systems, which include DGIS. Mr. Generous is the technical founder of DGIS along with Ms. Cotter as the conceptual founder.

Mr. William Spanos, Project Manager, CDC -- Mr. Spanos manages the contractor technical support staff, and as such maintains a selected complement that is highly motivated, intensely interested in the systems, and assures excellence in the continuity of the systems. The following people are or have been members of this critical staff, and have contributed to Hypermedia Laboratory activities in peripheral but important manners:

Mr. Rick Dunbar, Mr. Dan Gillilan, Ms. Nina Jones, Ms. Karen Kaye, Mr. Jon Krueger (joint author with Duc Tran on the X-Windows paper in this review), *Mr. Joe Morton, Mr. Andy Papp, Mr. Cliff Schenk.*

Effective 31 December 1988 the
Defense Applied Information Technology Center (DAITC)
was disestablished

Effective 1 January 1989 the
DAITC was replaced by the
Defense Technical Information Center Special Projects Office (DTICSPO).

The activities and management structure of DTICSPO remain the same.